

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 12:51:18 ; Search time 446.127 Seconds
(without alignments)
11200.072 Million cell updates/sec

Title: US-09-745-763-35
Perfect score: 1851
Sequence: 1 GGCTAGGCGCCGAGCTTAGT.....CTGAAAAAAAAAAAAAAAAA 1851

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03:.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1851	100.0	1851	19	AAV82779
2	1851	100.0	1851	24	ABO92016
3	1846.2	99.7	1863	21	AAZ98034
4	1846.2	99.7	1863	22	AAAD1647
5	1846.2	99.7	1863	24	ABK65743
6	1846.2	99.7	1863	25	ACC50817
7	1846.2	99.7	1863	25	ABZ71453
8	1833.4	99.0	1923	21	AAZ98139

9	1830.8	98.9	1884	21	AAZ58313	Human peptidase NA
10	1830.2	98.9	2077	25	ACC50603	Human secreted pro
11	1830.2	98.9	2077	25	ABZ71331	Secreted protein-e
12	1820.4	98.3	1860	22	AAK94491	Human full-length
13	1751.2	94.6	1767	21	AAK40493	Human fetal kidney
14	1751	94.6	1778	18	AAV02296	Human secreted pro
15	1617.2	87.4	1895	22	AAH9703	Human protein enco
16	1491	80.6	1784	23	AAH73592	DNA encoding novel
17	1145	61.9	1778	24	ABK62598	Rat sequence diffe
18	1088.4	59.3	1134	21	AAZ88101	Human secreted pro
19	1088.4	59.3	1134	22	AAAD1174	Human secreted pro
20	1088.4	59.3	1134	24	ABK69810	Human secreted pro
21	1088.4	59.3	1134	25	ACC50818	Human secreted pro
22	1088.4	59.3	1134	25	ABZ71454	Secreted protein-e
23	865.8	46.8	895	21	AAA44369	Human secreted exp
24	569	30.7	609	20	AAV87487	EST clone BU45. H
25	568.2	30.7	642	22	AAK92181	Human CDNA 5'-end
26	568.2	30.7	642	22	AAK93447	Human CDNA clone r
27	410.6	22.2	492	22	AAK92932	Human CDNA 3'-end
28	375.8	20.3	394	14	AAO61260	Human brain expres
29	375	20.3	424	23	AAZ73591	DNA encoding novel
30	324.4	17.5	360	18	AAH88078	3' portion of CDNA
31	324.4	17.5	360	18	AAV02198	Human secreted pro
32	324.4	17.5	360	18	AAV02140	Human secreted pro
33	310.2	16.8	317	20	AAK40587	Human secreted pro
34	290.8	15.7	314	20	AAV86369	EST clone AM282.
35	284.6	15.4	462	18	AAH97298	Human secreted pro
36	284.6	15.4	462	18	AAH88060	Partial CDNA clone
37	284.6	15.4	462	18	AAV02139	Human secreted pro
38	277.8	15.0	441	22	AAK56831	Human immune/haema
39	233	12.6	234	24	AAH97386	Gene #3884 used to
40	122.4	6.6	217	20	AAK40583	Human secreted pro
41	72	3.9	72	16	AAZ24970	Human gene signal
42	60.2	3.3	65	24	ABN33183	Mouse spliced tran
43	60	3.2	60	24	ABN36212	Human spliced tran
44	45.4	2.5	40862	24	ABU34073	Human immune syste
45	43	2.3	469	23	ABV56638	Human prostate exp

ALIGNMENTS

RESULT 1
AAV82779 standard; CDNA; 1851 BP.
AAV82779;
25-FEB-1999 (first entry)
Clone bu45_2 isolated from human adult placenta CDNA library.
Secreted protein; nutritional activity; immune stimulating; vaccine;
Suppressing activity; haematopoiesis regulating activity;
tissue growth activity; activin; inhibin activity; chemotaxis;
chemokine activity; haemostasis; thrombolytic activity; receptor;
ligand; anti-inflammatory; cadherin; tumour invasion suppressor;
tumour inhibition; gene therapy; ds.
Homo sapiens.
WO9842739-A2.
01-OCT-1998.
20-MAR-1998; 98MO-US05653.
19-MAR-1998; 98US-004446.
21-MAR-1997; 97US-0822167.
(GENY) GENETICS INST INC.
Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding V, Treacy M;
 XX MPI: 1998-609890/51.
 DR P-ESDB; AAM85456.
 XX
 PT New polynucleotides encoding secreted human proteins - derived from
 PT human foetal brain, adult brain, foetal kidney, placenta or adult
 PT pineal gland cDNA libraries.
 XX
 PS Claim 14; Page 69-70; 113pp; English.
 XX
 CC The present sequence encodes a secreted protein. The polynucleotide and
 CC secreted protein are predicted to have biological activities which would
 CC make them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is given.
 CC Suggested activities include nutritional activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity (no data is
 CC given in the specification to support these activities). The
 CC polynucleotide is also stated to be useful for gene therapy.
 CC
 SQ Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;

Query Match 100.0%; Score 1851; DB 19; Length 1851;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTAGGCGCGAGCTTAACTGCTCTGGAGCGCGCTCCGTGCGCGCGTAGAGCCGCCCTTA 60
 DB 1 GCGTAGGCGCGAGCTTAACTGCTCTGGAGCGCGCTCCGTGCGCGCGTAGAGCCGCCCTTA 60
 QY 61 TCAGATTAATCTTAAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTGCAT 120
 DB 61 TCAGATTAATCTTAAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTGCAT 120
 QY 121 TTTTGGGAGGTGACCTTTATCCCTGCTGCTGGGAAAGCTATATGCAAGAAATGCA 180
 DB 121 TTTTGGGAGGTGACCTTTATCCCTGCTGCTGGGAAAGCTATATGCAAGAAATGCA 180
 QY 181 TCTTAAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGGAGATGTTGCTA 240
 DB 181 TCTTAAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGGAGATGTTGCTA 240
 QY 181 TCTTAAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGGAGATGTTGCTA 240
 DB 181 TCTTAAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGGAGATGTTGCTA 240
 QY 241 AAGCAATATCACTAGCTGTTATGTTAAAGCCCAAGACAGATCTATAGCGATTGG 300
 DB 241 AAGCAATATCACTAGCTGTTATGTTAAAGCCCAAGACAGATCTATAGCGATTGG 300
 QY 301 CACTCTGCTGATATCTGTTGAGCCAGACTGAGTGGCTCCAAAGAACTTAAAGAAAGCA 360
 DB 301 CACTCTGCTGATATCTGTTGAGCCAGACTGAGTGGCTCCAAAGAACTTAAAGAAAGCA 360
 QY 361 TCCAAATTAATGACAAACCTGCAAGAGATGGCTGGAAGAAATTTCACTGAGCCAG 420
 DB 361 TCCAAATTAATGACAAACCTGCAAGAGATGGCTGGAAGAAATTTCACTGAGCCAG 420
 QY 421 TGAGAAATACCCCACTGGAGAGGAGAGAAATACAGCTGTGATGCTGAGCCAAATTC 480
 DB 421 TGAGAAATACCCCACTGGAGAGGAGAGAAATACAGCTGTGATGCTGAGCCAAATTC 480
 QY 481 ATTAATATGCCATCTCTGGGTCTTGGCAGAGAGATTTGGGACTCTCCAGAAAGCATTAAG 540
 DB 481 ATTAATATGCCATCTCTGGGTCTTGGCAGAGAGATTTGGGACTCTCCAGAAAGCATTAAG 540
 QY 541 CAGAGATCTGGGTGAGACTCTTTCGATGAACCTGAGAGAAAGGCTCAGAAAGCAAG 600
 DB 541 CAGAGATCTGGGTGAGACTCTTTCGATGAACCTGAGAGAAAGGCTCAGAAAGCAAG 600
 QY 601 GGAAGATTTGTTTATTAACAACCTTACATCACTCAAGGAGCGGTGCAATACCGAA 660
 DB 601 GGAAGATTTGTTTATTAACAACCTTACATCACTCAAGGAGCGGTGCAATACCGAA 660

QY 661 CGCAGGGGGGCGGTGGAAGCTGCGCAAGGTGGGGGCTTTGGCATCTCTCATTCGATCCGNG 720
 DB 661 CGCAGGGGGGCGGTGGAAGCTGCGCAAGGTGGGGGCTTTGGCATCTCTCATTCGATCCGNG 720
 QY 721 CCTCTTCTCCATATACAGTCTCTCACAAGATATTCAGAAATACAGAGATGCGTGGCCA 780
 DB 721 CCTCTTCTCCATATACAGTCTCTCACAAGATATTCAGAAATACAGAGATGCGTGGCCA 780
 QY 781 AAATTCACAGCCTGTATTAAGGTGGAAGATGCAAAATGATGTCAAGATGGCTTCTC 840
 DB 781 AAATTCACAGCCTGTATTAAGGTGGAAGATGCAAAATGATGTCAAGATGGCTTCTC 840
 QY 841 ATGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGACCTTACCCAGATCTGAT 900
 DB 841 ATGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGACCTTACCCAGATCTGAT 900
 QY 901 CCTTCAACACTGTATGACAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCA 960
 DB 901 CCTTCAACACTGTATGACAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCA 960
 QY 961 GTGACATCTGACAGCTGGAGATGTTGGCAGGGTGCATGATGATGGCGGTGAGACT 1020
 DB 961 GTGACATCTGACAGCTGGAGATGTTGGCAGGGTGCATGATGATGGCGGTGAGACT 1020
 QY 1021 TTAATATGAGGAAGCACTGCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTC 1080
 DB 1021 TTAATATGAGGAAGCACTGCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTC 1080
 QY 1081 TGCGGCTGTGCTCTGAGACTGCAAGAAACAGGTGAGTGTGCTTCCAGATTAATC 1140
 DB 1081 TGCGGCTGTGCTCTGAGACTGCAAGAAACAGGTGAGTGTGCTTCCAGATTAATC 1140
 QY 1141 AGTTACCAAGGTAAATATTTCCAACTACAGTCTGTGATGAGTCTGAGCGAGAACT 1200
 DB 1141 AGTTACCAAGGTAAATATTTCCAACTACAGTCTGTGATGAGTCTGAGCGAGAACT 1200
 QY 1201 TCTTACCACTGGGCTGCAATTCACCTGCGAGTAAAGGCGGAGGCGCATATGAGAGAG 1260
 DB 1201 TCTTACCACTGGGCTGCAATTCACCTGCGAGTAAAGGCGGAGGCGCATATGAGAGAG 1260
 QY 1261 TTAATGACCTGTGCAAGCCCTCAATATCACTCAGTCTGAGCCATGAGAGAGAGAG 1320
 DB 1261 TTAATGACCTGTGCAAGCCCTCAATATCACTCAGTCTGAGCCATGAGAGAGAGAG 1320
 QY 1321 ACATCAACCTTTGATTCAGAGCTGAGTGGTCTGAGAGCCATCTATGATGATTAACA 1380
 DB 1321 ACATCAACCTTTGATTCAGAGCTGAGTGGTCTGAGAGCCATCTATGATGATTAACA 1380
 QY 1381 AGTATTTCTTCTTCATCACTCCAGAGAGACACATGACATGATGATGATGATGATG 1440
 DB 1381 AGTATTTCTTCTTCATCACTCCAGAGAGACACATGACATGATGATGATGATGATG 1440
 QY 1441 TGAATGTTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTTGGCAGATGGAAGAAA 1500
 DB 1441 TGAATGTTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTTGGCAGATGGAAGAAA 1500
 QY 1501 TGTGCTAGGTCTTGAAGAAAGTAAAGAAAGAAAGCTTTTCACTGCTTGGCCAGATC 1560
 DB 1501 TGTGCTAGGTCTTGAAGAAAGTAAAGAAAGAAAGCTTTTCACTGCTTGGCCAGATC 1560
 QY 1561 CTGGGTCTGCAACTTTGAAAACTCCCTTCACTAAATTCATCAATTCATCTTCA 1620
 DB 1561 CTGGGTCTGCAACTTTGAAAACTCCCTTCACTAAATTCATCAATTCATCTTCA 1620
 QY 1621 AAGCAACCTATTTATGATCTTCTGTTATATCTTCTTGAATCACTTCCAAATTTCT 1680
 DB 1621 AAGCAACCTATTTATGATCTTCTGTTATATCTTCTTGAATCACTTCCAAATTTCT 1680
 QY 1681 GATTTGAAAAAAGAAATCAATTCCTCCCTTCCCAACATBAATCAATATGTA 1740
 DB 1681 GATTTGAAAAAAGAAATCAATTCCTCCCTTCCCAACATBAATCAATATGTA 1740


```

Db      661 CGCAGGCGCGGNGAGAGCTGCCAAGTGGGGCTTTGGCATCTCATTCAGTCCTGG 720
Qy      721 CCTCCTTCCTCACTACAGTCTCTCAACAGTATTCAGGAATACCGAGATGGCGGCCCA 780
Db      721 CCTCCTTCCTCACTACAGTCTCTCAACAGTATTCAGGAATACCGAGATGGCGGCCCA 780
Qy      781 AAATTCACAACAGCTGTATTTAGGTGGAAGATGCAGAAATGATGTCAGAAATGGCTTCTC 840
Db      781 AAATTCACAACAGCTGTATTTAGGTGGAAGATGCAGAAATGATGTCAGAAATGGCTTCTC 840
Qy      841 ATGGGATCAAAATTTGTCATTAGCTAAGATGGGGCAAGACTTACCCAGATCTGATT 900
Db      841 ATGGGATCAAAATTTGTCATTAGCTAAGATGGGGCAAGACTTACCCAGATCTGATT 900
Qy      901 CCTCAACAGCTGAGAGATCACTGGGAGCAAAATACAGAAAGGTTGATCTGCTCA 960
Db      901 CCTCAACAGCTGAGAGATCACTGGGAGCAAAATACAGAAAGGTTGATCTGCTCA 960
Qy      961 GTGCACTCTGGACAGCTGGAGATGTTGGCAGGGTGCCATGATGATGGCGGTGAGCCT 1020
Db      961 GTGCACTCTGGACAGCTGGAGATGTTGGCAGGGTGCCATGATGATGGCGGTGAGCCT 1020
Qy      1021 TTATATCATGGAAGACCTCTCACTTATTAAGATCTTGCGCTGCTCCAAAGAGACTC 1080
Db      1021 TTATATCATGGAAGACCTCTCACTTATTAAGATCTTGCGCTGCTCCAAAGAGACTC 1080
Qy      1081 TGGCGCTGGTGTCTGGAAGCTGGAAGAAACAAGTGGAGTGTGCTTCCAGTATATTC 1140
Db      1081 TGGCGCTGGTGTCTGGAAGCTGGAAGAAACAAGTGGAGTGTGCTTCCAGTATATTC 1140
Qy      1141 AGTTACACAGGTAATATTTTCCAACTACAGTCTGGTGAATGAGTCTGACGAGAACT 1200
Db      1141 AGTTACACAGGTAATATTTTCCAACTACAGTCTGGTGAATGAGTCTGACGAGAACT 1200
Qy      1201 TCTTACCCACTGGGCTGCAATTCATCTGGCAGTGAAGAGGCCATCATGAGAGAG 1260
Db      1201 TCTTACCCACTGGGCTGCAATTCATCTGGCAGTGAAGAGGCCATCATGAGAGAG 1260
Qy      1261 TTATAGGCTGCTGAGAGCCCTCAATATGACTCAGTCTGAGCCATGGAAGAGGACAG 1320
Db      1261 TTATAGGCTGCTGAGAGCCCTCAATATGACTCAGTCTGAGCCATGGAAGAGGACAG 1320
Qy      1321 ACATCAACTTTTGAATCCAGCTGAGAGTCTGAGCCAGTCTGATGATGACTTATACA 1380
Db      1321 ACATCAACTTTTGAATCCAGCTGAGAGTCTGAGCCAGTCTGATGATGACTTATACA 1380
Qy      1381 AGTATTTCTTCTTCATCACTCTCCACGAGACACATGATGATCCAAAGCAGA 1440
Db      1381 AGTATTTCTTCTTCATCACTCTCCACGAGACACATGATGATCCAAAGCAGA 1440
Qy      1441 TGAATGTCGTCGCTGTTGGGCTGTTTCTTATGTTGTCGACATGGAAGAA 1500
Db      1441 TGAATGTCGTCGCTGTTGGGCTGTTTCTTATGTTGTCGACATGGAAGAA 1500
Qy      1501 TGTCTGCTAGGTCTCTAGAAACAGTAAGAAAGAAAGCTTTTCATGCTTGGCAGAAATC 1560
Db      1501 TGTCTGCTAGGTCTCTAGAAACAGTAAGAAAGAAAGCTTTTCATGCTTGGCAGAAATC 1560
Qy      1561 CTGGGCTGCGAATTTGGAAAATCTCCTCTTCACATTAACAATTCATCTTCA 1620
Db      1561 CTGGGCTGCGAATTTGGAAAATCTCCTCTTCACATTAACAATTCATCTTCA 1620
Qy      1621 AAGCAACTCTATTTCAATGCTTCTGTTATATCTTTCTTGATCTTCCAAATCTCT 1680
Db      1621 AAGCAACTCTATTTCAATGCTTCTGTTATATCTTTCTTGATCTTCCAAATCTCT 1680
Qy      1681 GATTCTAGAAAAGAAATCAATCTCCCTCCCTCCCAACAATAGATCAATATGTTGA 1740
Db      1681 GATTCTAGAAAAGAAATCAATCTCCCTCCCTCCCAACAATAGATCAATATGTTGA 1740
Qy      1741 GGGATTAACAGTGGGGCAATTTCTTATATCACTCTTATTAACATGTTTCCACTTTAA 1800
Db      1741 GGGATTAACAGTGGGGCAATTTCTTATATCACTCTTATTAACATGTTTCCACTTTAA 1800

```

```

Qy      1801 AGTAACACTTAATTAATTTTGAAGATCTCGAAGAAAAA 1851
Db      1801 AGTAACACTTAATTAATTTTGAAGATCTCGAAGAAAAA 1851

RESULT 3
AA298034
ID AA298034 standard; cDNA, 1863 BP.
XX
AC AA298034;
XX
DT 09-MAY-2000 (first entry)
XX
DE Human secreted protein encoding nucleotide sequence SEQ ID NO:28.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiinflammatory; neurotropic; neuroprotective; antiallergic; cancer;
KW tumour; neurodegenerative disorder; developmental abnormality; allergy;
KW foetal deficiency; blood disorder; immune system disorder; arthritis;
KW autoimmune disease; hepatic disease; renal disease; inflammation;
KW Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
KW infection; AIDS; spinal cord injury; transplant rejection; diabetes;
KW asthma; sepsis; acne; psoriasis; cardiovascular disorder;
KW reproductive disorder; gastrointestinal disorder; respiratory disorder;
KW metabolic disorder; food additive; preservative; ss.

OS Homo sapiens.
PN WO200004140-A1.
PD 27-JAN-2000.
PF 14-JUL-1999; 99WO-US15849.
PR 15-JUL-1998; 98US-0092921.
PR 15-JUL-1998; 98US-0092922.
PR 15-JUL-1998; 98US-0092956.
PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
PI Lafleur DW, Ehner R, Olsen HS, Brewer LA, Florence KA, Young PE;
PI Muceneki M, Endress GA, Soppet DR;
XX
XX WPI: 2000-161128/14.
DR P-PSDB; AAY87081.
XX
XX New isolated human genes, useful for diagnosis and treatment of, e.g.
XX cancers, neurological or blood disorders -
XX
XX Claim 1; Page 319; 494pp; English.
XX
XX The polynucleotide sequences given in AA298017 to AA298108 encode the
XX human secreted proteins given in AAY87064 to AAY87223. Human secreted
XX protein can have activities based on the tissues and cells the genes are
XX expressed in. Examples of activities include: cytostatic;
XX immunosuppressive; antiinflammatory; neurotropic; neuroprotective; and
XX antiallergic. The polynucleotides and their corresponding secreted
XX polypeptides are useful for preventing, treating or ameliorating medical
XX conditions, e.g. by protein or gene therapy. Also pathological conditions
XX can be diagnosed by determining the amount of the new polypeptides in a
XX sample or by determining the presence of mutations in the new
XX polynucleotides. Human secreted protein s and their polynucleotides can
XX be used for developing products for the diagnosis or treatment of cancer,
XX tumours, neurodegenerative disorders, developmental abnormalities and
XX foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, hepatic and renal disease, inflammation,
XX allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
XX osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
XX cardiovascular disorders, reproductive disorders, gastrointestinal
XX disorders, respiratory disorders and metabolic disorders. The

```

CC proteins or polynucleotides can also be used as food additives or
CC preservatives. The proteins are also useful for identifying their
CC binding partners. AA298008 to AA298016 and AA298063 are sequence used in
CC the exemplification of the present invention.

XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 99.7%; Score 1846.2; DB 21; Length 1863;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 GGCTAGGCGCGGAGCTTATGCTCTGAGAGCGCGCTCGCGCGCTGAGAGCGCGCTTA 60
DB 1 GACTAGGCGCGGAGCTTATGCTCTGAGAGCGCGCTCGCGCGCTGAGAGCGCGCTTA 60
QY 61 TCAGATTTATCTTAAACAGAAAAACCACTGAAAAAAATTAATTCCTTATCTTCGAT 120
DB 61 TCAGATTTATCTTAAACAGAAAAACCACTGAAAAAAATTAATTCCTTATCTTCGAT 120
QY 121 TTTTCGGTGTGTACCTTTATCCCTGTGCTGAGAAAGCTATATGCAAGATGCA 180
DB 121 TTTTCGGTGTGTACCTTTATCCCTGTGCTGAGAAAGCTATATGCAAGATGCA 180
QY 181 TCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTGCTA 240
DB 181 TCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTGCTA 240
QY 241 AAGCAATCATCACTTACTGCTTTATGTGTAAGCCCGAACAATCTTATGAGGATGG 300
DB 241 AAGCAATCATCACTTACTGCTTTATGTGTAAGCCCGAACAATCTTATGAGGATGG 300
QY 301 CACTTCTGTGTATACCTGTGTAAGCCCGAACAATCTTATGAGGATGG 360
DB 301 CACTTCTGTGTATACCTGTGTAAGCCCGAACAATCTTATGAGGATGG 360
QY 361 TCCTAAATTTATGTACCAAACTGTGAGCAAGATGGGCTGAGAAAGTTTCACTGAGCCAG 420
DB 361 TCCTAAATTTATGTACCAAACTGTGAGCAAGATGGGCTGAGAAAGTTTCACTGAGCCAG 420
QY 421 TGAAGATATCCCACTGTGAGAGGAGGAGAAATACAGCTGTGATGTGAGAGCAAGATTC 480
DB 421 TGAAGATATCCCACTGTGAGAGGAGGAGAAATACAGCTGTGATGTGAGAGCAAGATTC 480
QY 481 ATAAGATATCCCACTGTGAGAGGAGGAGAAATACAGCTGTGATGTGAGAGCAAGATTC 540
DB 481 ATAAGATATCCCACTGTGAGAGGAGGAGAAATACAGCTGTGATGTGAGAGCAAGATTC 540
QY 541 CAGAAGTTCTGTGTGACTCTTTTCGATGTAAGCTGCAAGAGAGGCGCTCAGAGCAAGAG 600
DB 541 CAGAAGTTCTGTGTGACTCTTTTCGATGTAAGCTGCAAGAGAGGCGCTCAGAGCAAGAG 600
QY 601 GGAAGATTTGTTTATTAACCAACTTACATCACTCAAGGAGCGGTGCAATTCGAA 660
DB 601 GGAAGATTTGTTTATTAACCAACTTACATCACTCAAGGAGCGGTGCAATTCGAA 660
QY 661 CGCAGGCGGCGGTGAGAGCTGCAAGAGGAGGCGGTGAGAGCTGCAATTCGAGTCCGAG 720
DB 661 CGCAGGCGGCGGTGAGAGCTGCAAGAGGAGGCGGTGAGAGCTGCAATTCGAGTCCGAG 720
QY 721 CCTCTCTTCATCTACAGTCTCTCACAAGATTCAGAGATTCAGAGATTCAGAGATTC 780
DB 721 CCTCTCTTCATCTACAGTCTCTCACAAGATTCAGAGATTCAGAGATTCAGAGATTC 780
QY 781 AAATTCACAAGCTGTATTAAGGTGAGAGATGCAAGATTAAGTCAAGATTCGCTTTC 840
DB 781 AAATTCACAAGCTGTATTAAGGTGAGAGATGCAAGATTAAGTCAAGATTCGCTTTC 840
QY 841 ATGGATCAAAATTTGTCAATCACTGTAAGATGAGGAGCAAGATTCAGATTCGATTC 900
DB 841 ATGGATCAAAATTTGTCAATCACTGTAAGATGAGGAGCAAGATTCAGATTCGATTC 900
QY 901 CCTTCACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGACTGTCA 960

```

```

DB 901 CCTTCACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGACTGTCA 960
QY 961 GTGAGATCTGAGCAGCTGGAGATGTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
DB 961 GTGAGATCTGAGCAGCTGGAGATGTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
QY 1021 TTATATCATGGAGAGCAGCTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGAGCTC 1080
DB 1021 TTATATCATGGAGAGCAGCTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGAGCTC 1080
QY 1081 TGGGCTGTGCTGTGAGCTGCAAGAGAGCAAGAGTGTGAGTGTGCTTCCAGTATATC 1140
DB 1081 TGGGCTGTGCTGTGAGCTGCAAGAGAGCAAGAGTGTGAGTGTGCTTCCAGTATATC 1140
QY 1141 AGTTACACAAGATTAATTTTCCAACTACAGCTGTGATGAGGAGTGTGAGGAGGAGGAGG 1200
DB 1141 AGTTACACAAGATTAATTTTCCAACTACAGCTGTGATGAGGAGTGTGAGGAGGAGGAGG 1200
QY 1201 TCTTACCACTGGGCTGCAATTCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
DB 1201 TCTTACCACTGGGCTGCAATTCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
QY 1261 TTATGAGCTGTGAGCAGCTCTCAATATCACTAGTCTGTGAGCAGTGAAGAGGAGCAG 1320
DB 1261 TTATGAGCTGTGAGCAGCTCTCAATATCACTAGTCTGTGAGCAGTGAAGAGGAGCAG 1320
QY 1321 ACATCAACTTTTGAATCCAGCTGAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
DB 1321 ACATCAACTTTTGAATCCAGCTGAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
QY 1381 AGTATTTCTTCTTCATCACTCCACGAGAGACACATGATGATGATGATGATGATGATGATG 1440
DB 1381 AGTATTTCTTCTTCATCACTCCACGAGAGACACATGATGATGATGATGATGATGATGATG 1440
QY 1441 TGAATGTTGCTGTGCTGTGTTGGGCTGTGTTTCTTATGTTGTTGAGACATGAGAGAA 1500
DB 1441 TGAATGTTGCTGTGCTGTGTTGGGCTGTGTTTCTTATGTTGTTGAGACATGAGAGAA 1500
QY 1501 TGTGCTTGTGCTGTGCTGTGTTGGGCTGTGTTTCTTATGTTGTTGAGACATGAGAGATC 1560
DB 1501 TGTGCTTGTGCTGTGCTGTGTTGGGCTGTGTTTCTTATGTTGTTGAGACATGAGAGATC 1560
QY 1561 CTGGGCTGTGCACTTTGAGAACTCTCTTCAATCAATCAATCAATCAATCAATCAATCA 1620
DB 1561 CTGGGCTGTGCACTTTGAGAACTCTCTTCAATCAATCAATCAATCAATCAATCAATCA 1620
QY 1621 AAGCAACTCTATTTATGCTGTGTTTATTTATTTATTTATTTATTTATTTATTTATTTCT 1680
DB 1621 AAGCAACTCTATTTATGCTGTGTTTATTTATTTATTTATTTATTTATTTATTTATTTCT 1680
QY 1681 GATTTTGAAGAAAGAAATCAATTCCTCCCTCCCTCCACACATGAGATCAATATGATA 1740
DB 1681 GATTTTGAAGAAAGAAATCAATTCCTCCCTCCCTCCACACATGAGATCAATATGATA 1740
QY 1741 GGAATTAAGTGGGAGATTTCTTATATCACTCTTAAAGAACTGTGTTTCACTTTAAA 1800
DB 1741 GGAATTAAGTGGGAGATTTCTTATATCACTCTTAAAGAACTGTGTTTCACTTTAAA 1800
QY 1801 AGTAAACACTTAATTAATTTTGGAGATCTGTGAAAAAATTTTAAAAAATTTTAAAAA 1851
DB 1801 AGTAAACACTTAATTAATTTTGGAGATCTGTGAAAAAATTTTAAAAAATTTTAAAAA 1851

```

RESULT 4
AD11647
ID AD11647 standard; cDNA; 1863 BP.
XX AD11647;
XX AC
XX 24-SEP-2001 (first entry)
XX Human secreted protein-encoding gene 18 cDNA clone HRAU35, SEQ ID NO:28.
XX

KM Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 KM foetal abnormality; developmental abnormality; haematopoietic disorder;
 KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KM Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KM psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KM inflammation; neurological disorder; Alzheimer's disease; food additive;
 KM angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 KM pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KM cell culture; chemotaxis; vulnerary; binding partner identification;
 KM gene therapy; ss.
 KM Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FT CDS 99..1517
 FT /*tag= a
 FT /product= "Human secreted protein precursor"
 FT sig_peptide 99..170
 FT /*tag= b
 FT mat_peptide 171..1514
 FT /*tag= c
 FT /product= "Mature human secreted protein"
 FT
 FT
 XX WO200151504-A1.
 PN 19-JUN-2001.
 PD 12-JAN-2001; 2001WO-US00911.
 XX 13-JAN-2000; 2000US-0482273.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Komatoullis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Entress GA, Muscenski M, Ebner R;
 XX MPI; 2001-425865/45.
 DR P-PSDB; AAB06058.
 XX
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 FT used in preventing, treating or ameliorating a medical condition -
 XX
 PS Claim 1; Page 686; 864pp; English.
 XX AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
 CC protein genes, and AAB06041-AAB06132 represent the proteins they encode.
 CC AAB06133-AAB06205 represent human secreted protein fragments.
 CC The secreted proteins and their genes are useful for preventing, treating
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 CC Pathological conditions can be diagnosed by determining the amount of the
 CC new protein in a sample or by determining the presence of mutations in
 CC the new genes. Specific uses are described for each of the 71 genes,
 CC based on the tissues in which they are most highly expressed, and include
 CC developing products for the diagnosis or treatment of proliferative
 CC disorders, cancer, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
 CC pregnancy-related disorders, endocrine disorders, and infections. The
 CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, for supporting cell culture of primary tissues,
 CC to regenerate tissues, to identify their cognate ligands or binding
 CC partners, and in chemotaxis, and can be used as a food additive or
 CC preservative to modify storage properties. Antibodies specific for a
 CC protein of the invention can be used in alleviating symptoms associated
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
 CC The present sequence represents a human secreted protein-encoding cDNA of

CC the invention.
 XX
 XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
 Query Match 99.7%; Score 1846.2; DB 22; Length 1863;
 Best Local Similarly 99.8%; Pred. No. 0;
 Matches 1848; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 1 GGCTAGGCGCGGACCTTAGTCCTGGAGCGCGCTCCGTCGCGCGCCGACAGCGCCCTTA 60
 1 GACTAGGCGCGGACCTTAGTCCTGGAGCGCGCTCCGTCGCGCGCCGACAGCGCCCTTA 60
 61 TCAGATTATCTTAAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTCGAT 120
 61 TCAGATTATCTTAAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTCGAT 120
 121 TTTTCGGTGTGTTACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGATGGCA 180
 121 TTTTCGGTGTGTTACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGATGGCA 180
 181 TCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTGCTA 240
 181 TCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTGCTA 240
 241 AAGCAATCATCAACCTAGCTGTTTATGATAAGCCAGAACAGATCCTATGAGCATTTG 300
 241 AAGCAATCATCAACCTAGCTGTTTATGATAAGCCAGAACAGATCCTATGAGCATTTG 300
 301 CACTTCTGTGTTAGTACTGTTTGAAGCCAGACTGAGTGTCTCCAAAGACCTAAGAAAGCA 360
 301 CACTTCTGTGTTAGTACTGTTTGAAGCCAGACTGAGTGTCTCCAAAGACCTAAGAAAGCA 360
 361 TCCAAATTATGATCAAAACCTGCGAGCAAGATGGGCTGGAAGAAATTCACCTGGAGCCAG 420
 361 TCCAAATTATGATCAAAACCTGCGAGCAAGATGGGCTGGAAGAAATTCACCTGGAGCCAG 420
 421 TGAGAAATACCCCACTGGAGAGGGAGAGAAATCAGCTGTGATCTGGAGCCAAAGATTC 480
 421 TGAGAAATACCCCACTGGAGAGGGAGAGAAATCAGCTGTGATCTGGAGCCAAAGATTC 480
 481 ATTAAGATGACCATCTCTGGGTTTGGCAGACGATTTGGGACTCTCCCAAGGCAATTACG 540
 481 ATTAAGATGACCATCTCTGGGTTTGGCAGACGATTTGGGACTCTCCCAAGGCAATTACG 540
 541 CAGAAGTTCTGAGTGTGACCTCTTTCGATGAATGACAGAGAGGGCCCTCAGAACCAAGAG 600
 541 CAGAAGTTCTGAGTGTGACCTCTTTCGATGAATGACAGAGAGGGCCCTCAGAACCAAGAG 600
 601 GGAAGATTGTTGTTATTAACCACTTACATCAACTCAAGAGCGGTGCAATACCGAA 660
 601 GGAAGATTGTTGTTATTAACCACTTACATCAACTCAAGAGCGGTGCAATACCGAA 660
 661 CGCAGGGGGCGGTGAGAGCTCCAAAGTGGGGCTTTGGCATCTCTCAATTCGATCCGGG 720
 661 CGCAGGGGGCGGTGAGAGCTCCAAAGTGGGGCTTTGGCATCTCTCAATTCGATCCGGG 720
 721 CCTCTTCTCATCATCAAGCTCTCACAAGGATTTACGAATATCAGGATGAGCGTGGCCA 780
 721 CCTCTTCTCATCATCAAGCTCTCACAAGGATTTACGAATATCAGGATGAGCGTGGCCA 780
 781 AATTTCCAAACGCTGTATTAACGTTGAGAGATGCAAGATATGATCAAGATGGCTTTC 840
 781 AATTTCCAAACGCTGTATTAACGTTGAGAGATGCAAGATATGATCAAGATGGCTTTC 840
 841 ATGGGATCAAAATTTGCTATTAGCTAAAGATGGGGGCAACCACTCAACCAATCTGAT 900
 841 ATGGGATCAAAATTTGCTATTAGCTAAAGATGGGGGCAACCACTCAACCAATCTGAT 900
 901 CTTTCAACACTGTGAGAGAGATCACTGGGAGCAAAATATCCAGAACAGTTGATCTGTCA 960
 901 CTTTCAACACTGTGAGAGAGATCACTGGGAGCAAAATATCCAGAACAGTTGATCTGTCA 960
 961 GTGACATCTGAGAGCTGGAGATGTTGGGAGGGGTGCATGATGATGCGGTGGAGCCT 1020

```

Db      961 GTGGACATCTGGACAGCTGGGATGTTGGGCGGGTGCATGATGAGCGGTGAGACCT 1020
Qy      1021 TTATATATGAGGAGACCTCTCACTTATTAAGATCTTGGCTGCGCTCAAAAGGAGCTC 1080
Db      1021 TTATATATGAGGAGACCTCTCACTTATTAAGATCTTGGCTGCGCTCAAAAGGAGCTC 1080
Qy      1081 TGGGCTGCTGCTCTGAGCTGAGAGAACAGAGGTGAGTGGCTTCCAGTATATATC 1140
Db      1081 TGGGCTGCTGCTCTGAGCTGAGAGAACAGAGGTGAGTGGCTTCCAGTATATATC 1140
Qy      1141 AGTTACCAAGTAATATTTTCCAACTACAGTCTGGTATGAGTCTGACGACAGAACT 1200
Db      1141 AGTTACCAAGTAATATTTTCCAACTACAGTCTGGTATGAGTCTGACGACAGAACT 1200
Qy      1201 TCTTACCCACTGGGCTGCAATTCACCTGAGTGAAGGCGGAGGCAATCATGAGAGG 1260
Db      1201 TCTTACCCACTGGGCTGCAATTCACCTGAGTGAAGGCGGAGGCAATCATGAGAGG 1260
Qy      1261 TTATGAGCCTGCTGACGCCCCCTCAATATCACTCAGTCTGAGCGCATGAGAGAGGAG 1320
Db      1261 TTATGAGCCTGCTGACGCCCCCTCAATATCACTCAGTCTGAGCGCATGAGAGAGGAG 1320
Qy      1321 ACATCAACTTTTGGATCCAGCTGAGTGGAGTGGAGCCAGTCTAATTGATGACTTATACA 1380
Db      1321 ACATCAACTTTTGGATCCAGCTGAGTGGAGTGGAGCCAGTCTAATTGATGACTTATACA 1380
Qy      1381 AGATTTCTCTCTCCATCATCTCCACGAGACACCATGATGATGATGATGATGATGATGAT 1440
Db      1381 AGATTTCTCTCTCTCCATCATCTCCACGAGACACCATGATGATGATGATGATGATGATGAT 1440
Qy      1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db      1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Qy      1501 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db      1501 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Qy      1561 CTGGGCTGCTGCACTTTGGAGAAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1620
Db      1561 CTGGGCTGCTGCACTTTGGAGAAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1620
Qy      1621 AAGCAACCTATTTGATGCTTCTGTTATCTTCTGATGATCTTCTGATGATCTTCTGATGAT 1680
Db      1621 AAGCAACCTATTTGATGCTTCTGTTATCTTCTGATGATCTTCTGATGATCTTCTGATGAT 1680
Qy      1681 GATTTGAGAGAGAGATCATCTCCCTCCCTCCCAACATGAGATCAATATGATGATGAT 1740
Db      1681 GATTTGAGAGAGAGATCATCTCCCTCCCTCCCAACATGAGATCAATATGATGATGAT 1740
Qy      1741 GGGATTAACAGTGGGCGCATTTCTTATATACCTCTTAAACATGTTTCCACTTTAA 1800
Db      1741 GGGATTAACAGTGGGCGCATTTCTTATATACCTCTTAAACATGTTTCCACTTTAA 1800
Qy      1801 AGTAACACCTTAATAATTTTGGAGAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1851
Db      1801 AGTAACACCTTAATAATTTTGGAGAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1851

```

```

KW      virucide; fungicide; ophthalmological; autoimmune disease; neoplasm;
KW      rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KW      cardiovascular disorder; cerebrovascular disorder; cerebral ischemia;
KW      angiodysplasia; nervous system disorder; Alzheimer's disease; infection;
KW      ocular disorder; corneal infection; wound healing; skin aging;
KW      epithelial cell proliferation; food additive.
XX
XX      Homo sapiens.
XX
XX      NC0200226931-A2.
XX
XX      04-APR-2002.
XX
XX      24-SEP-2001; 2001MO-US29871.
XX
XX      25-SEP-2000; 2000US-234925P.
XX      12-JAN-2001; 2001MO-US05911.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;
XX      Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
XX      Endreess GA, Mucenski M, Edner R;
XX
XX      WPI; 2002-362489/39.
XX      P-PSDB; ABG33880.
XX
XX      Novel 71 isolated secreted polypeptides and polynucleotides encoding
XX      the polypeptides, useful for treating Huntington's disease, sepsis,
XX      meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
XX      asthma
XX
XX      Claim 1; Page 1170; 1478pp; English.
XX
XX      The invention relates to an isolated nucleic acid molecule (or its
XX      fragment, homologue complement or allelic variant) encoding a human
XX      secreted protein (and its fragment, domain, epitope, variant, secreted
XX      form and species variant). Also included are a recombinant vector
XX      comprising the nucleic acid, a recombinant host cell comprising the
XX      vector, an antibody against the secreted protein, a recombinant host cell
XX      that expresses the secreted protein and a method of identifying a binding
XX      partner of the secreted protein. The nucleic acid and protein are used to
XX      prevent, diagnose, treat or ameliorate a medical condition in e.g.
XX      humans, mice, rabbits, goats, cats, dogs, chickens or sheep
XX      for example autoimmune diseases e.g. rheumatoid arthritis,
XX      hyperproliferative disorders e.g. neoplasms of the breast or liver,
XX      cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
XX      e.g. cerebral ischemia, angiodysplasia, nervous system disorders e.g.
XX      Alzheimer's disease, infections caused by bacteria, viruses and fungi and
XX      ocular disorders e.g. corneal infection. Many other diseases and
XX      disorders are listed in the specification. The polypeptides can also be
XX      used to aid wound healing an epithelial cell proliferation, to prevent
XX      skin aging due to sunburn, to maintain organs before transplantation, for
XX      supporting cell culture of primary tissues, to regenerate tissues and in
XX      chemotaxis. The polypeptides can also be used as a food additive or
XX      preservative to increase or decrease storage capabilities. The present
XX      sequence encodes a novel human secreted protein of the invention.
XX
XX      Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
XX
XX      Query Match      99.7%; Score 1846.2; DB 24; Length 1863;
XX      Best Local Similarity 99.8%; Pred. No. 0;
XX      Matches 1848; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy      1 GGCTAGGCGCGAGCTTATGCTTGGAGCGGCTCGCTGCGCGCGCTGAGCGCGCCTTA 60
Db      1 GACTAGGCGCGAGCTTATGCTTGGAGCGGCGGCTCGCTGCGCGCGCTGAGCGCGCCTTA 60
Qy      61 TCAGATTATCTTAAACAGAGAAACCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db      61 TCAGATTATCTTAAACAGAGAAACCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy      121 TTTTGGTGTGTTTCACTTTATCCCTGCTCTGGAGAGAGATATATGAGAGATGGA 180

```

```

Db 121 TTTTGGTGTGTCACTTTATCCCTGTGCTGTGGAAAGCTATATGCAAGATGGCA 180
Qy 181 TCTCTAAGAGACTTTTGAAGAAATPAAAGAAATAGCCAGCTGTGAGATGTTGCTA 240
Db 181 TCTCTAAGAGACTTTTGAAGAAATPAAAGAAATAGCCAGCTGTGAGATGTTGCTA 240
Qy 241 AAGCATCATCAACCTAGCTGTTTATGTGTAAGCCAGAAAGATCCTATGAGCGATTGG 300
Db 241 AAGCATCATCAACCTAGCTGTTTATGTGTAAGCCAGAAAGATCCTATGAGCGATTGG 300
Qy 301 CACTTCTGGTGTATCTGTTGAGCCGAGCTGAGTGGCTTCAAGAACTTGAAGAAAGCCA 360
Db 301 CACTTCTGGTGTATCTGTTGAGCCGAGCTGAGTGGCTTCAAGAACTTGAAGAAAGCCA 360
Qy 361 TCCAAATTAATGACAAACCTGAGCAAGATGGCTGAGAAAGTTCCCTGGAGCCAG 420
Db 361 TCCAAATTAATGACAAACCTGAGCAAGATGGCTGAGAAAGTTCCCTGGAGCCAG 420
Qy 421 TGAGAAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGTGATGCTGAGCCAGAAATTC 480
Db 421 TGAGAAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGTGATGCTGAGCCAGAAATTC 480
Qy 481 ATAGATAGCCATCTGAGGTCTGTGGCAGCAGCATTGGGACTCTCCAGAAAGCATTACAG 540
Db 481 ATAGATAGCCATCTGAGGTCTGTGGCAGCAGCATTGGGACTCTCCAGAAAGCATTACAG 540
Qy 541 CAGAAATCTGTGTGTGATCTCTTTCGATGAACTGAGAGAAAGGCTTCAAGAGCAAGAG 600
Db 541 CAGAAATCTGTGTGTGATCTCTTTCGATGAACTGAGAGAAAGGCTTCAAGAGCAAGAG 600
Qy 541 CAGAAATCTGTGTGTGATCTCTTTCGATGAACTGAGAGAAAGGCTTCAAGAGCAAGAG 600
Db 541 CAGAAATCTGTGTGTGATCTCTTTCGATGAACTGAGAGAAAGGCTTCAAGAGCAAGAG 600
Qy 601 GGAAGATGTTGTTTATTAACCAACCTTACATCACTACTCAAGAGAGCGGTGCAATACGGA 660
Db 601 GGAAGATGTTGTTTATTAACCAACCTTACATCACTACTCAAGAGAGCGGTGCAATACGGA 660
Qy 661 CGCAGAGGCGGTGAGAGCTGCAAGGTGGGCGCTTTGGCATCTCTCATTCGATCCGTGG 720
Db 661 CGCAGAGGCGGTGAGAGCTGCAAGGTGGGCGCTTTGGCATCTCTCATTCGATCCGTGG 720
Qy 721 CCTCTCTTCTCCATCTACAGTCTCTCAACAGGTATTCAGGAATACAGAGATGGCGCCCA 780
Db 721 CCTCTCTTCTCCATCTACAGTCTCTCAACAGGTATTCAGGAATACAGAGATGGCGCCCA 780
Qy 781 AAATTCCAAACAGCTGTATATAGGTGAGAGATGCAAGAAATGATGCAAGATGGCTCTC 840
Db 781 AAATTCCAAACAGCTGTATATAGGTGAGAGATGCAAGAAATGATGCAAGATGGCTCTC 840
Qy 841 ATGGATCAAAATTTGTCAATTCAGTAAAGATGGGCGCAAAAGACTACCAAGATCTGAT 900
Db 841 ATGGATCAAAATTTGTCAATTCAGTAAAGATGGGCGCAAAAGACTACCAAGATCTGAT 900
Qy 901 CCTTCAACAGCTGTACAGAGATCACTGGAGCAAAATTCACAAACAGGTGTGACGTGCA 960
Db 901 CCTTCAACAGCTGTACAGAGATCACTGGAGCAAAATTCACAAACAGGTGTGACGTGCA 960
Qy 961 GTGACATCTGAGACAGCTGAGATGTTGGCAGAGTGGCCATGAGATGGCGGTGAGAGCT 1020
Db 961 GTGACATCTGAGACAGCTGAGATGTTGGCAGAGTGGCCATGAGATGGCGGTGAGAGCT 1020
Qy 1021 TTAATATCATGGAGACACTCTCATTTAAAGATCTTGGGCTGGCTCAAGAGAGACTC 1080
Db 1021 TTAATATCATGGAGACACTCTCATTTAAAGATCTTGGGCTGGCTCAAGAGAGACTC 1080
Qy 1081 TGGCGCTGTGTCTGTGAGCTGAGAAAGAAAGAGGTGAGTGTGCTTCCAGTATTTATC 1140
Db 1081 TGGCGCTGTGTCTGTGAGCTGAGAAAGAAAGAGGTGAGTGTGCTTCCAGTATTTATC 1140
Qy 1141 AGTTACACAGGTAAATATTTCCAACTACAGCTGTGATGAGAGTCTGACGAGAGAACT 1200
Db 1141 AGTTACACAGGTAAATATTTCCAACTACAGCTGTGATGAGAGTCTGACGAGAGAACT 1200
Qy 1201 TCTTACCACTGGGCTGCAATTCATGAGAGTGAAGAGCCAGGCGCATGAGAGAGG 1260

```

```

Db 1201 TCTTACCACTGGGCTGCAATTCATGAGAGTGAAGAGCCAGGCGCATGAGAGAGG 1260
Qy 1261 TTAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 TTAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy 1321 ACATCACTTTTGGATCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 ACATCACTTTTGGATCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 AGATATTTCTTCCATCACTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 AGATATTTCTTCCATCACTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Qy 1501 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 1501 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Qy 1561 CTGGGCTGCAACTTTGAGAAAGCTCTTCAACATTAATTCATCAATTCATTCATTCAT 1620
Db 1561 CTGGGCTGCAACTTTGAGAAAGCTCTTCAACATTAATTCATCAATTCATTCATTCAT 1620
Qy 1621 AAGCAACATCTATTTGATGCTTCTGTTATATCTTGTGATGATCTTGTGATGATCTTGT 1680
Db 1621 AAGCAACATCTATTTGATGCTTCTGTTATATCTTGTGATGATCTTGTGATGATCTTGT 1680
Qy 1681 GATTCAGAAAGAGAGATCACTTCCCTCCCTCCCAACATGAGATCAACATATGATA 1740
Db 1681 GATTCAGAAAGAGAGATCACTTCCCTCCCTCCCAACATGAGATCAACATATGATA 1740
Qy 1741 GGAATTAAGTGGGAGATTTCTTATATCACTCTTAAACATGTTGTTCACTTTAAA 1800
Db 1741 GGAATTAAGTGGGAGATTTCTTATATCACTCTTAAACATGTTGTTCACTTTAAA 1800
Qy 1801 AGTAAACATTAATTAATTTTGAAGATCTCTGAAAAAATTTTGAAGATCTCTGAAAAA 1851
Db 1801 AGTAAACATTAATTAATTTTGAAGATCTCTGAAAAAATTTTGAAGATCTCTGAAAAA 1851

RESULT 6
ACC50817
ID ACC50817 standard; cDNA; 1863 BP.
XX
AC ACC50817;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted protein coding sequence, SEQ ID 484.
XX
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;
XX antiinflammatory; nootropic; neuroprotective;
XX antiparkinsonian; gene therapy; human; cardiovascular disorder;
XX gene; ss.
XX
OS Homo sapiens.
XX
PN MO200295010-A2.
XX
PD 28-NOV-2002.
XX
PF 19-MAR-2002; 2002WO-US09785.
XX
PR 21-MAR-2001; 2001US-27340P.
XX 19-JUN-2001; 2001US-306171P.
XX 13-NOV-2001; 2001US-331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;

```


XX WPI; 2003-129429/12.
DR Novel human secreted proteins, useful for detecting, preventing,
XX diagnosing, prognosticating, treating and/or ameliorating
PT cardiovascular disorders such as arrhythmia -
XX
XX Claim 21; SEQ ID 484; 1881bp; English.
XX
XX The present invention relates to novel human secreted proteins
CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The
CC proteins and their coding sequences are useful for the preparation of a
CC diagnostic or pharmaceutical composition for diagnosing or treating a
CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
CC coronary arteriosclerosis and myocardial ischemia), neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, renal disorders,
CC proliferative disorders and/or cancerous diseases and conditions, for
CC wound healing and epithelial cell proliferation, to treat inflammation or
CC infection, for treating thrombosis and arteriosclerosis, for treating or
CC preventing neural damage which occurs in neuronal disorders or
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
CC disease, to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts, to prevent skin aging or hair loss, to
CC stimulate growth and differentiation of haematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism.
CC Note: The sequence data for this patent was published in electronic
CC format and is available from WIPO at
CC fep.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
Query Match 99.7%; Score 1846.2; DB 25; Length 1863;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGCTAGGCGCGGAGCTTATGCTCTGGAGCCGCTCGTGGCGGCTGAGAGCCGCTTA 60
Db 1 GACTAGGCGCGGAGCTTATGCTCTGGAGCCGCTCGTGGCGGCTGAGAGCCGCTTA 60
QY 61 TCAGATTATCTTAAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCCGAT 120
Db 61 TCAGATTATCTTAAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCCGAT 120
QY 121 TTTTGGGGGTGCTCACTTTTATCCCTGCTGCTGGGAAGCTATATGCAAGATGGCA 180
Db 121 TTTTGGGGGTGCTCACTTTTATCCCTGCTGCTGGGAAGCTATATGCAAGATGGCA 180
QY 181 TCTCTAAGAGAGCTTTGGAAGAAATGAAGAAATAGCAGCTGTGAGATTTGCTA 240
Db 181 TCTCTAAGAGAGCTTTGGAAGAAATGAAGAAATAGCAGCTGTGAGATTTGCTA 240
QY 241 AAGCAATCATCACTAGCTGTTATGTTAAAGCCAGACAGATCTTATAGAGATTGG 300
Db 241 AAGCAATCATCACTAGCTGTTATGTTAAAGCCAGACAGATCTTATAGAGATTGG 300
QY 301 CACTTCGTGTTGATCTGTGAGCCAGACTGAGTGGTCCAGAGAACTAGAAAAGCA 360
Db 301 CACTTCGTGTTGATCTGTGAGCCAGACTGAGTGGTCCAGAGAACTAGAAAAGCA 360
QY 361 TCCAAATATATGACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAGCCAG 420
Db 361 TCCAAATATATGACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAGCCAG 420
QY 421 TGAGATATCCCACTGGAGAGGGAGAAAGATCACTGCTGTGATGCTGAGCCAGAAATTC 480
Db 421 TGAGATATCCCACTGGAGAGGGAGAAAGATCACTGCTGTGATGCTGAGCCAGAAATTC 480
QY 481 ATAAGATAGCATCTCGGTCTTGGCAGACAGATTGGGACTCTCCAGAGGCAATTACAG 540

Db 481 ATAAGATAGCATCTCGGTCTTGGCAGACAGATTGGGACTCTCCAGAGGCAATTACAG 540
QY 541 CAGAGCTTGGTGGTGGAGCTCTTGTGATGAACTGAGAGAGGGCTTCAAGAGCAAG 600
Db 541 CAGAGCTTGGTGGTGGAGCTCTTGTGATGAACTGAGAGAGGGCTTCAAGAGCAAG 600
QY 601 GGAAGATGTTGTTTAAACCAACCTTACATCACTACTCAAGAGCGTGCATACCGAA 660
Db 601 GGAAGATGTTGTTTAAACCAACCTTACATCACTACTCAAGAGCGTGCATACCGAA 660
QY 661 CGAGGGGGGGTGAAGCTGCCAGGTGGGGGCTTTGCAATCTCATTCATTCGTCGG 720
Db 661 CGAGGGGGGGTGAAGCTGCCAGGTGGGGGCTTTGCAATCTCATTCATTCGTCGG 720
QY 721 CTTCTTCTTCATCTACAGCTCTCAACAGATATTCAGAAATCCAGATGGGCTGCCA 780
Db 721 CTTCTTCTTCATCTACAGCTCTCAACAGATATTCAGAAATCCAGATGGGCTGCCA 780
QY 781 AAATTCCAACAGCTGATTAAGGTGAGATGCAAGAAATGATGTCAGAAATGGCTTTC 840
Db 781 AAATTCCAACAGCTGATTAAGGTGAGATGCAAGAAATGATGTCAGAAATGGCTTTC 840
QY 841 ATGGATCAAAATTTGATTCATTCAGCTTAAAGATGGGGGCAAGACCTACCGATACGATT 900
Db 841 ATGGATCAAAATTTGATTCATTCAGCTTAAAGATGGGGGCAAGACCTACCGATACGATT 900
QY 901 CTTTCAACCTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAACAGTTGTACTGTCA 960
Db 901 CTTTCAACCTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAACAGTTGTACTGTCA 960
QY 961 GTGAGACATCGAGAGTGGAGTGGTGGGAGGGTGCATGATGATGGGGTGGAGCT 1020
Db 961 GTGAGACATCGAGAGTGGAGTGGTGGGAGGGTGCATGATGATGGGGTGGAGCT 1020
QY 1021 TTATATCATGGGAAGACCTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTC 1080
Db 1021 TTATATCATGGGAAGACCTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTC 1080
QY 1081 TGGGGCTGTGCTCTGAGACTGCAAGAGAACAGAGTGGAGTTGGCTTCACTATATTC 1140
Db 1081 TGGGGCTGTGCTCTGAGACTGCAAGAGAACAGAGTGGAGTTGGCTTCACTATATTC 1140
QY 1141 AGTTACCAAGATTAATTTTCAACTACAGCTGTGGATGAGTGCAGCGAGAACT 1200
Db 1141 AGTTACCAAGATTAATTTTCAACTACAGCTGTGGATGAGTGCAGCGAGAACT 1200
QY 1201 TCTTACCACTGGGGCTGCAATTCAGTGGCAGTGAAGAGCCAGGGCCATCATGAGAGG 1260
Db 1201 TCTTACCACTGGGGCTGCAATTCAGTGGCAGTGAAGAGCCAGGGCCATCATGAGAGG 1260
QY 1261 TTATGAGCTGTGAGAGCCCTTCAATATCACTCACTGCTGAGCCATGAGAGAGGAG 1320
Db 1261 TTATGAGCTGTGAGAGCCCTTCAATATCACTCACTGCTGAGCCATGAGAGAGGAG 1320
QY 1321 ACATCAACTTTTGAATTCAGAGTGGAGTGGCTGAGAGCACTTGAAGCTTATACA 1380
Db 1321 ACATCAACTTTTGAATTCAGAGTGGAGTGGCTGAGAGCACTTGAAGCTTATACA 1380
QY 1381 AGTATTTCTTCTTCATCACTCCAGAGAGACACATGATGATGATCAAGAGCA 1440
Db 1381 AGTATTTCTTCTTCATCACTCCAGAGAGACACATGATGATGATCAAGAGCA 1440
QY 1441 TGAATGTTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTTGGCAGACATGGAAGAA 1500
Db 1441 TGAATGTTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTTGGCAGACATGGAAGAA 1500
QY 1501 TGTGCTAGAGTCTTAAACAGATTAAGAAAGAAAGCTTTCAGTCTGCGCAGAGATC 1560
Db 1501 TGTGCTAGAGTCTTAAACAGATTAAGAAAGAAAGCTTTCAGTCTGCGCAGAGATC 1560
QY 1561 CTGGGCTGCAACTTTGAAAACCTCTTCACTAATCAATTCATCAATTCATCA 1620

Db 1561 CTGGGCTGCAACTTTGGAAAACTCCTCTTCACATACATTTTCATTCATCTTCA 1620
 Qy 1621 AAGCAACTCTATTTTCTGCTTTCTGTATATTTCTTGATCTTCCAAATTTCTT 1680
 Db 1621 AAGCAACTCTATTTTCTGCTTTCTGTATATTTCTTGATCTTCCAAATTTCTT 1680
 Qy 1681 GATTTCTGAAAAAGAAATCATTTCTCCCTCCCTCCACACATAGATCAATATGTA 1740
 Db 1681 GATTTCTGAAAAAGAAATCATTTCTCCCTCCCTCCACACATAGATCAATATGTA 1740
 Qy 1741 GGAATTACAGTGGGGGCAATTTCTTATATCACTTAAATAAATTTTCCATTAA 1800
 Db 1741 GGAATTACAGTGGGGGCAATTTCTTATATCACTTAAATAAATTTTCCATTAA 1800
 Qy 1801 AGTAACTCTTATATTTTGGAAATCTCTGAAAAAATTTTCAATTTTCA 1860
 Db 1801 AGTAACTCTTATATTTTGGAAATCTCTGAAAAAATTTTCAATTTTCA 1860
 RESULT 7
 AB271453
 ID AB271453 standard; cDNA; 1863 BP.
 AC AB271453;
 XX
 DT 04-APR-2003 (first entry)
 XX
 DE Secreted protein-encoding gene 142 cDNA clone HRAC35, SEQ ID NO:274.
 XX
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
 KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
 KW antiinflammatory; immunosuppressive; vulnery; chromosome 8q22.2;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200276488-A1.
 XX
 PD 03-OCT-2002.
 XX
 PF 19-MAR-2002; 2002MO-US08276.
 XX
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-029900/02.
 DR P-PSDB; ABR00274.
 XX
 FT New human secreted proteins and nucleic acids, useful for detecting,
 FT preventing, diagnosing, prognosticating, treating and/or ameliorating
 FT e.g. gastrointestinal diseases and disorders, or cancers -
 XX
 PS Claim 21; Page 918; 1216p; English.
 XX
 CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary

CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein-
 CC encoding cDNA clone of the invention.
 XX
 SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
 Query Match 99.7%; Score 1846.2; DB 25; Length 1863;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1848; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 GGCTAGCCCGAGCCTTAGTCTGAGAGCCGCTCCGTCGAGAGCCGCCCTTA 60
 Db 1 GACTAGCCCGAGCCTTAGTCTGAGAGCCGCTCCGTCGAGAGCCGCCCTTA 60
 Qy 61 TCAGATTATCTTAAACAAGAAACCACTGAGAAAAAATGAAATTCCTTATCTTGCAT 120
 Db 61 TCAGATTATCTTAAACAAGAAACCACTGAGAAAAAATGAAATTCCTTATCTTGCAT 120
 Qy 121 TTTTCGGTGGTGTACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGATGCA 180
 Db 121 TTTTCGGTGGTGTACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGATGCA 180
 Qy 181 TCTCTAAGAGACCTTTTGAAGAAATTAAGAAATATGCCAGCTGTGAGATGTGTA 240
 Db 181 TCTCTAAGAGACCTTTTGAAGAAATTAAGAAATATGCCAGCTGTGAGATGTGTA 240
 Qy 241 AAGAAATCATCACTGCTGTTATGTTAAGCCCAAGACAGATCTATGAGGATTTG 300
 Db 241 AAGAAATCATCACTGCTGTTATGTTAAGCCCAAGACAGATCTATGAGGATTTG 300
 Qy 301 CACTTCTGTATATCTTGTGACCCAGACTGATGCTCCAGAACTTAAAGAAAGCA 360
 Db 301 CACTTCTGTATATCTTGTGACCCAGACTGATGCTCCAGAACTTAAAGAAAGCA 360
 Qy 361 TCCAATTTATGTATACCAAAACCTGACAGAGATGGCTGAGAAAGTTTCACTGAGCCAG 420
 Db 361 TCCAATTTATGTATACCAAAACCTGACAGAGATGGCTGAGAAAGTTTCACTGAGCCAG 420
 Qy 421 TGAAATATCCCACTGGAGAGGGGAGAAAGATAGCTGTGATCTGAGAGCAAGATTTC 480
 Db 421 TGAAATATCCCACTGGAGAGGGGAGAAAGATAGCTGTGATCTGAGAGCAAGATTTC 480
 Qy 481 ATAGATATGCACTCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAATTACG 540
 Db 481 ATAGATATGCACTCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAATTACG 540
 Qy 541 CAGAAATTTCTGTGTGACCTCTTTGATGAATGTCAGAGAGAGGCTCAGAGCAAGAG 600
 Db 541 CAGAAATTTCTGTGTGACCTCTTTGATGAATGTCAGAGAGAGGCTCAGAGCAAGAG 600
 Qy 601 GGAAGATTTGTTTATTAACAACCTTACATCACTCTCAAGAGCGTGTGAATACGAA 660
 Db 601 GGAAGATTTGTTTATTAACAACCTTACATCACTCTCAAGAGCGTGTGAATACGAA 660
 Qy 661 CGCAGGGGCGGTGGAAGCTCCAAAGTGGGGCTTTGGCATCTCTCATTTGCATCCGTGG 720
 Db 661 CGCAGGGGCGGTGGAAGCTCCAAAGTGGGGCTTTGGCATCTCTCATTTGCATCCGTGG 720
 Qy 721 CTTCTTCTCATCTTACAGTCTTCAACAGGATTTACAGAAATACAGAGATGGCTGCCA 780
 Db 721 CTTCTTCTCATCTTACAGTCTTCAACAGGATTTACAGAAATACAGAGATGGCTGCCA 780
 Qy 781 AAATTCACAGCCTGATATTCGCTGAGAAAGATGCAAAATATGCAAGATTCCTTCTC 840
 Db 781 AAATTCACAGCCTGATATTCGCTGAGAAAGATGCAAAATATGCAAGATTCCTTCTC 840
 Qy 841 ATGGATCAAAATTTGATTCAGCTAAAGATGGGGGCAAGACCTTACCCAGATCTGAT 900


```

Oy 1 GGGTAGGCGGAGCTTAGTCTCTGGGAGCCGCTCGTCGCCGCCGTAGAGCGCCCTTA 60
Db 81 GGGTAGGCGGAGCTTAGTCTCTGGGAGCCGCTCGTCGCCGCCGTAGAGCGCCCTTA 140
Oy 61 TCAGATTTATCTTTACAGAGAAAACAAGTGGAAAAAATGAAATTCCTTATCTTCGAT 120
Db 142 TCAGATTTATCTTTACAGAGAAAACAAGTGGAAAAAATGAAATTCCTTATCTTCGAT 200
Oy 121 TTTTGGGTTGGTTTCACTTTTATCCCTGTGCTCTGGGAAAGACTATATGCAAGATGGCA 180
Db 201 TTTTGGGTTGGTTTCACTTTTATCCCTGTGCTCTGGGAAAGACTATATGCAAGATGGCA 260
Oy 181 TCTTAAGAGGACTTTTGAAGAAATAAAAGAAATAGCCAGCTGTGGAGATGTTGCTA 240
Db 261 TCTTAAGAGGACTTTTGAAGAAATAAAAGAAATAGCCAGCTGTGGAGATGTTGCTA 320
Oy 241 AAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCTATAGCGATTGG 300
Db 321 AAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCTATAGCGATTGG 380
Oy 301 CACTTCTGGTTGATCTGTTGGAACCAAGTGGCTCCAGAACTTGAAGAAAGCCA 360
Db 381 CACTTCTGGTTGATCTGTTGGAACCAAGTGGCTCCAGAACTTGAAGAAAGCCA 440
Oy 361 TCCAAATTTATGTACCAAAAACCTGACGAAAGATGGGCTGGAGAAAGTTGACCTGGAGCGAG 420
Db 441 TCCAAATTTATGTACCAAAAACCTGACGAAAGATGGGCTGGAGAAAGTTGACCTGGAGCGAG 500
Oy 421 TGAGAAATACCCCACTGGGAGAGAGGAGAAATCAAGCTGTGATGCTGAGAGCCAGAAATTC 480
Db 501 TGAGAAATACCCCACTGGGAGAGAGGAGAAATCAAGCTGTGATGCTGAGAGCCAGAAATTC 560
Oy 481 ATAAATATGCCATCTCGGGTCTTTGGAGAGAGATTTGGACTCTCCAGAAAGCATTTACAG 540
Db 561 ATAAATATGCCATCTCGGGTCTTTGGAGAGAGATTTGGACTCTCCAGAAAGCATTTACAG 620
Oy 541 CAGAGTTCTGGTGGTGAACCTCTTCGATGAACTGACAGAGAGAGGCTCAGAGAGAGAG 600
Db 621 CAGAGTTCTGGTGGTGAACCTCTTCGATGAACTGACAGAGAGAGGCTCAGAGAGAGAG 680
Oy 601 GGAAGATTTGTTTATTAACAACCTTACATCACTCAAGAGAGAGAGTGAATACCA 660
Db 681 GGAAGATTTGTTTATTAACAACCTTACATCACTCAAGAGAGAGAGTGAATACCA 740
Oy 661 CGCAGGCGGCGGTGGAAGCTGCCAAGGTGGGCTTTGGCATCTCATTCGATCCGTGG 720
Db 741 CGCAGGCGGCGGTGGAAGCTGCCAAGGTGGGCTTTGGCATCTCATTCGATCCGTGG 800
Oy 721 CCTCCTTCCTCACTACAGTCTCTCAACAGGATTCAGGAAATACAGAGATGGCGGCCA 780
Db 801 CCTCCTTCCTCACTACAGTCTCTCAACAGGATTCAGGAAATACAGAGATGGCGGCCA 860
Oy 781 AAATTCACAAGCCGTGTATTAAGGTGGAAGATGACAGAAATGATGTCAAGATGGCTTCTC 840
Db 861 AGATTCACAAGCCGTGTATTAAGGTGGAAGATGACAGAAATGATGTCAAGATGGCTTCTC 920
Oy 841 ATGGATCAAAATTTGTCAATTCAGCTAAAGATGGGCGCAAAAGCTTACCCAGATCTGATT 900
Db 921 ATGGATCAAAATTTGTCAATTCAGCTAAAGATGGGCGCAAAAGCTTACCCAGATCTGATT 980
Oy 901 CCTTCAACACCTGTACAGAGATCACTGGGAGCAAAATTCAGAAACAGGTTGTACTGGTCA 960
Db 981 CCTTCAACACCTGTGTACAGAGATCACTGGGAGCAAAATTCAGAAACAGGTTGTACTGGTCA 1040
Oy 961 GTGACATCTGGAGAGCTGGAGATGTTGGGCAAGGCTGATGATGAGCGGTGGAGCCT 1020
Db 1041 GTGACATCTGGAGAGCTGGAGATGTTGGGCAAGGCTGATGATGAGCGGTGGAGCCT 1100
Oy 1021 TTATATCATGGAGAGACTTCACTTATTAAGATCTTGGGCTGGCTCAAAAGAGAGACTC 1080
Db 1101 TTATATCATGGAGAGACTTCACTTATTAAGATCTTGGGCTGGCTCAAAAGAGAGACTC 1160

```

```

Oy 1081 TGCGGCTGTGCTCTGTGACCTGACAGAGAAACAAGGTGAGATTGTGCTTCCAGATTATTC 1140
Db 1161 TGCGGCTGTGCTCTGTGACCTGACAGAGAAACAAGGTGAGATTGTGCTTCCAGATTATTC 1220
Oy 1141 AGTTACACAGGTAATATTTTCCAACTACAGTCTGTGTGATGAGATCTGACCGCAGAACCT 1200
Db 1221 AGTTACACAGGTAATATTTTCCAACTACAGTCTGTGTGATGAGATCTGACCGCAGAACCT 1280
Oy 1201 TCTTACCACTGGGCTGCAATTCAGTGGCAGGTGAAAGGCGGAGCCATCATGGAGGAGG 1260
Db 1281 TCTTACCACTGGGCTGCAATTCAGTGGCAGGTGAAAGGCGGAGCCATCATGGAGGAGG 1340
Oy 1261 TTATGAGCTGTGCTGACGCCCCCTCAATATCACTCAGGTCTTGAGCCATGAGAGAGGACG 1320
Db 1341 TTATGAGCTGTGCTGACGCCCCCTCAATATCACTCAGGTCTTGAGCCATGAGAGAGGACG 1400
Oy 1321 ACATCAACTTTTGGATCCAGCTGAGAGTGCCTGGAGCCAGTCTACTTGAATGATATACA 1380
Db 1401 ACATCAACTTTTGGATCCAGCTGAGAGTGCCTGGAGCCAGTCTACTTGAATGATATACA 1460
Oy 1381 AGTATTTCTTCTTCATCACTCCGACGAGACACATGACATGATGGAATCCAAAGCAGA 1440
Db 1461 AGTATTTCTTCTTCATCACTCCGACGAGACACATGACATGATGGAATCCAAAGCAGA 1500
Oy 1441 TGAATGTTGCTGCTGCTGCTGTTGGGCTGTGTTTCTTATGTTGTGACAGACATGAGAGAA 1500
Db 1521 TGAATGTTGCTGCTGCTGCTGTTGGGCTGTGTTTCTTATGTTGTGACAGACATGAGAGAA 1580
Oy 1501 TGCTGCTGAGGCTCTAGAGAAACGATGAGAGAGAAAGGTTTCATGCTTGGCCAGGATTC 1560
Db 1581 TGCTGCTGAGGCTCTAGAGAAACGATGAGAGAGAAAGGTTTCATGCTTGGCCAGGATTC 1640
Oy 1561 CTGGGCTGAGAACTTTGAAATCTCTTCACTCAATCAATTCATCAATTCATCTTCA 1620
Db 1641 CTGGGCTGAGAACTTTGAAATCTCTTCACTCAATCAATTCATCAATTCATCTTCA 1700
Oy 1621 AAGCAAACTTATTTCAATGCTTCTGTATTTATCTTCTGATACCTTCCAAATTCCTCT 1680
Db 1701 AAGCAAACTTATTTCAATGCTTCTGTATTTATCTTCTGATACCTTCCAAATTCCTCT 1760
Oy 1681 GATTCTGAGAAAAAGAAATCATTTCTCCCTCCCTCCACACATGAAATCAATATGCTTA 1740
Db 1761 GATTCTGAGAAAAAGAAATCATTTCTCCCTCCCTCCACACATGAAATCAATATGCTTA 1820
Oy 1741 GGGATTCAGTGGGCGGCTTTCTTATATACCTCTTAAATAACATGTTTCCACTTAAA 1800
Db 1821 GGGATTCAGTGGGCGGCTTTCTTATATACCTCTTAAATAACATGTTTCCACTTAAA 1880
Oy 1801 AGTAAACACTTAATTAATTTTGGAGAGTCTGTAAGAAAAA 1843
Db 1881 AGTAAACACTTAATTAATTTTGGAGAGTCAAAAAA 1923

RESULT 9
AAZ58313
ID AAZ58313 standard; cDNA; 1884 BP.
XX
AC AAZ58313;
XX
DT 08-MAY-2000 (first entry)
XX
DE Human peptidase NAALAD-ase IV cDNA.
XX
KW NAALAD-ase IV; N-acetylated alpha-linked acidic dipeptidase; human;
KW Chromosome 8q21.3; prostate cancer; neurodegenerative disease;
KW Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
KW peripheral neuropathy; Huntington's disease; acute brain injury;
KW multiple sclerosis; peripheral nerve trauma; ischemia; dementia;
KW gene therapy; diagnosis; nootropic; neuroprotective; neuroleptic;
KW antiparkinsonian; anticonvulsant; vasotropic; ss.
XX
OS Homo sapiens.
XX

```

FH Key Location/Qualifiers
 FT CDS 149..1567
 FT /*tag= a
 XX MO200004157-A2.
 XX 27-JAN-2000.
 XX 14-JUL-1999; 99WO-GB02241.
 XX 14-JUL-1998; 98GB-0015284.
 XX (JANC) JANSSEN PHARM NV.
 XX Pangalos M, Neefs JEFM, Peeters DCG;
 XX WPI: 2000-182424/16.
 DR P-PSDB; AAY58879.
 XX
 PT New human N-acetylated alpha-linked acidic dipeptidases for treating
 PT neural disorders e.g. Alzheimer's disease, schizophrenia and
 PT Parkinson's disease
 XX
 PS Claim 7, Fig 5; 95pp; English.
 XX
 CC The present sequence is that of cDNA coding for human
 CC N-acetylated alpha-linked acidic dipeptidase IV (NALAD-ase IV,
 CC see AAY58879). The cDNA was obtained from a gall bladder cDNA
 CC library. Analysis of the open reading frame predicts a type II
 CC integral membrane protein with 5 potential N-glycosylation sites.
 CC The NALAD-ase II gene was mapped to chromosome 8q21.3.
 CC NALAD-ase IV expression was low in all tissues examined by RT-PCR.
 CC The invention provides human NALAD-ase I, II and IV polypeptides,
 CC cDNAs, antisense nucleic acids, vectors, host cells, transgenic
 CC organisms, antagonists and agonists. These are useful for treating
 CC neural disorders such as Alzheimer's disease, schizophrenia, ALS,
 CC Parkinson's disease, peripheral neuropathy, Huntington's disease,
 CC acute brain injury, multiple sclerosis, exposure to neurotoxins,
 CC peripheral nerve trauma, ischemia or dementia (claimed). Nucleic
 CC acids can also be used for gene therapy and for genetic screening
 CC of predisposition to disorders associated with NALAD-ase.
 XX
 SQ Sequence 1884 BP; 520 A; 432 C; 458 G; 474 T; 0 other;
 Query Match 98.9%; Score 1830.8; DB 21; Length 1884;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1832; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 411 TCCTAATTTATGATACCAAACTCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAGCCAG 470
 Qy 421 TGAATAATACCCCTGGGAGAGGGAGAAATATAGCTGTGATCTGTGAGCAAGATTC 480
 Db 471 TGAGAAATACCCCACTGGAGAGGGAGAAATATAGCTGTGATCTGTGAGCAAGATTC 530
 Qy 481 ATAGATAGCCATCTGGGTCTTGGAGAGCATTCCTCCAGAAAGGCAATTACAG 540
 Db 531 ATAGATAGCCATCTGGGTCTTGGAGAGCATTCCTCCAGAAAGGCAATTACAG 590
 Qy 541 CAGAGTTCTGTGTGATGACCTCTTGTGATGAATGACAGAGAGGGCTCAGAGCAAG 600
 Db 591 CAGAGTTCTGTGTGATGACCTCTTGTGATGAATGACAGAGAGGGCTCAGAGCAAG 650
 Qy 601 GGAAGATTTGTTGTTTAAACCACTTACATCACTCTCAGAGAGCGGTGCAATACGGA 660
 Db 651 GGAAGATTTGTTGTTTAAACCACTTACATCACTCTCAGAGAGCGGTGCAATACGGA 710
 Qy 661 CGCAGGGGGGCGGTGAGAGCTGCCAAGGTGGGGCTTGGCATCTCTGATTCGATCCGTGG 720
 Db 711 CGCAGGGGGGCGGTGAGAGCTGCCAAGGTGGGGCTTGGCATCTCTGATTCGATCCGTGG 770
 Qy 721 CCTCTTCTCCATCTTACAGTCTTACACAGATATTCAGAAATACAGATGCGGTGCCA 780
 Db 771 CCTCTTCTCCATCTTACAGTCTTACACAGATATTCAGAAATACAGATGCGGTGCCA 830
 Qy 781 AATATCCAAAGCTGTATTAAGGTGAGAAATGATGCAAAATGATGCAAGATGCTTCTC 840
 Db 831 AGATTTCCAAAGCTGTATTAAGGTGAGAAATGATGCAAAATGATGCAAGATGCTTCTC 890
 Qy 841 ATGGAGTCAAAATTTGTATTCAGTAAAGATGGGGGCAAAAGCTTACCCAGATCTGATT 900
 Db 891 ATGGAGTCAAAATTTGTATTCAGTAAAGATGGGGGCAAAAGCTTACCCAGATCTGATT 950
 Qy 901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAGATGCTTACTGTCTCA 960
 Db 951 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAGATGCTTACTGTCTCA 1010
 Qy 961 GTGACATCTGAGACAGCTGGGATTTGGCAGAGGTGCCATGATGATGGCGGTGAGCCT 1020
 Db 1011 GTGACATCTGAGACAGCTGGGATTTGGCAGAGGTGCCATGATGATGGCGGTGAGCCT 1070
 Qy 1021 TTATATCATGGGAAGCACTCACTTATTAAGATCTTGGGCTGGCTCAAGAGGACTC 1080
 Db 1071 TTATATCATGGGAAGCACTCACTTATTAAGATCTTGGGCTGGCTCAAGAGGACTC 1130
 Qy 1081 TGCAGCTGTGCTCTGTGACTGCAAGAAACAAGGTGGAGTGGCTTCCAGATTTATC 1140
 Db 1131 TGCAGCTGTGCTCTGTGACTGCAAGAAACAAGGTGGAGTGGCTTCCAGATTTATC 1190
 Qy 1141 AGTTACCAAAAGTAAATATTTCCACTACAGTCTGTGATGAGTCTGAGCGCAGAACT 1200
 Db 1191 AGTTACCAAAAGTAAATATTTCCACTACAGTCTGTGATGAGTCTGAGCGCAGAACT 1250
 Qy 1201 TCTTACCCACTGGGCTGCAATTCACCTGGCAGTGAAGAAGCGCAGGCATCATGAGAGAG 1260
 Db 1251 TCTTACCCACTGGGCTGCAATTCACCTGGCAGTGAAGAAGCGCAGGCATCATGAGAGAG 1310
 Qy 1261 TTATGAGCTGTGCTGCAAGCTTCAATATCACTCAGTCTGAGCGATGAGAGAGGAGCAG 1320
 Db 1311 TTATGAGCTGTGCTGCAAGCTTCAATATCACTCAGTCTGAGCGATGAGAGAGGAGCAG 1370
 Qy 1321 ACATCAAACTTTTGGATCCAAAGTGAAGTCCCTGGAGCGCAGTCTTATGATGATCTTATCA 1380
 Db 1371 ACATCAAACTTTTGGATCCAAAGTGAAGTCCCTGGAGCGCAGTCTTATGATGATCTTATCA 1430
 Qy 1381 AGTATTTCTTCTTCATCACTCCAGAGGAGACATGATGATGATGATGATGATGATGATGAT 1440
 Db 1431 AGTATTTCTTCTTCATCACTCCAGAGGAGACATGATGATGATGATGATGATGATGATGAT 1490
 Qy 1441 TGAATGTGCTGTGCTGTGTTGGGCTGTTTCTTATGTTGTTGACAGATGAGAGAA 1500


```

Db      754 CCTCTCTCCATCTACAGTCTCCACAGAGTATTACAGAAATACAGAGATGGCTGCCA 813
Qy      761 AATTCGAACGCGTATTTACGGTGAAGATGCAAGAAATGTCAGAAATGGCTTCTC 840
Db      814 AGATTCGAACGCGTATTTACGGTGAAGATGCAAGAAATGTCAGAAATGGCTTCTC 873
Qy      841 ATGGGATCAAAATGTCATTCAGCTTAAAGATGGGGGCAAAAGCTACCCAGATCTGATT 900
Db      874 ATGGGATCAAAATGTCATTCAGCTTAAAGATGGGGGCAAAAGCTACCCAGATCTGATT 933
Qy      901 CCTTCAACTGTGACAGAGATCTGAGAGCAAAATTCAGAAACAGTTGACTGTCTCA 960
Db      934 CCTTCAACTGTGACAGAGATCTGAGAGCAAAATTCAGAAACAGTTGACTGTCTCA 993
Qy      961 GTGACATCTGGAACAGCTGGGATGTTGGGCGAGGCTGCCATGAGATGATGGCGTGAAGCT 1020
Db      994 GTGACATCTGGAACAGCTGGGATGTTGGGCGAGGCTGCCATGAGATGATGGCGTGAAGCT 1053
Qy      1021 TTATATCATGGGAAGCACTCTCATTTATTAAGATCTTGGGCTGCGTCCAAAGAGAGCTC 1080
Db      1054 TTATATCATGGGAAGCACTCTCATTTATTAAGATCTTGGGCTGCGTCCAAAGAGAGCTC 1113
Qy      1081 TGCGGCTGTGCTCTGGAAGCTGCAAGAAAGAGGTTGAGTGTGCTTCCAGATTATTC 1140
Db      1114 TGCGGCTGTGCTCTGGAAGCTGCAAGAAAGAGGTTGAGTGTGCTTCCAGATTATTC 1173
Qy      1141 AGTTACCAAGGTAATTTTCCAACTTCTGTCATGATGAGATCTGACCCAGAACT 1200
Db      1174 AGTTACCAAGGTAATTTTCCAACTTCTGTCATGATGAGATCTGACCCAGAACT 1233
Qy      1201 TCTTACCACTGGGCTGCAATTCCTGGCAGTGAAGAGGCGCCATCATGAGAGAG 1260
Db      1234 TCTTACCACTGGGCTGCAATTCCTGGCAGTGAAGAGGCGCCATCATGAGAGAG 1293
Qy      1261 TTATGAGCTGCTGCAAGCCCTCAATATCACTGACGCTGATGAGCATGAGAGAGAG 1320
Db      1294 TTATGAGCTGCTGCAAGCCCTCAATATCACTGACGCTGATGAGCATGAGAGAGAG 1353
Qy      1321 ACATCACTTTTGGATCCAACTGGAAGTGGCTGGAGCCAGTCTTATGATATTTATCA 1380
Db      1354 ACATCACTTTTGGATCCAACTGGAAGTGGCTGGAGCCAGTCTTATGATATTTATCA 1413
Qy      1381 AGTATTTCTTTCATCACTCCACGGAACACCATGACTGTATGATGATCCAAAGCAGA 1440
Db      1414 AGTATTTCTTTCATCACTCCACGGAACACCATGACTGTATGATGATCCAAAGCAGA 1473
Qy      1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTCAGACATGGAAGAA 1500
Db      1474 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTCAGACATGGAAGAA 1533
Qy      1501 TGTGCTCCTAGTCTGAGAAACAGTAAAGAAAGAGCTTTCATGCTTGGCCAGGAATC 1560
Db      1534 TGTGCTCCTAGTCTGAGAAACAGTAAAGAAAGAGCTTTCATGCTTGGCCAGGAATC 1593
Qy      1561 CTGGGCTGCAACTTTGGAAGAACTCCTTCAATTAATTAATTCATCCATTCATCTCA 1620
Db      1594 CTGGGCTGCAACTTTGGAAGAACTCCTTCAATTAATTAATTCATCCATTCATCTCA 1653
Qy      1621 AAGCACACTCTATTTCAATGCTTCTGTTATTTATCTTCTGATATCTTCCAAATTCCT 1680
Db      1654 AAGCACACTCTATTTCAATGCTTCTGTTATTTATCTTCTGATATCTTCCAAATTCCT 1713
Qy      1681 GATTCTAGAAAAAGAAATCATCTTCCCTCCCTCCACCAATGAAATCAATATGAGTA 1740
Db      1714 GATTCTAGAAAAAGAAATCATCTTCCCTCCCTCCACCAATGAAATCAATATGAGTA 1773
Qy      1741 GGGATTAAGTGGGGGCAATTTCTTATATCACTTAAATTAATGTTTCCACTTTAA 1800
Db      1774 GGGATTAAGTGGGGGCAATTTCTTATATCACTTAAATTAATGTTTCCACTTTAA 1833
Qy      1801 AGTAAACATTAATTAATTTTGAAGATCTCTGA 1835

```

```

Db      1834 AGTAAACATTAATTAATTTTGAAGATCTCTGA 1868
RESULT 11
AB271331
ID AB271331 standard; cDNA; 2077 BP.
XX
AC AB271331;
XX
DT 04-APR-2003 (first entry)
XX
DE Secreted protein-encoding gene 142 cDNA clone HRA035, SEQ ID NO:152.
XX
KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
KW antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN W020027648-A1.
XX
PD 03-OCT-2002.
XX
PF 19-MAR-2002; 2002MO-US08276.
XX
PR 21-MAR-2001; 2001US-277340P.
PR 19-JUL-2001; 2001US-306171P.
PR 13-NOV-2001; 2001US-331287P.
XX
PA (HUMA-) HUMAN GENOME SCT INC.
XX
PI Rosen CA, Ruben SM,
XX
WP1: 2003-029900/02.
DR P-PSDB: ABR00152.
XX
XX
PT New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers -
XX
XX
PS Claim 21, Page 841-842; 1216pp; English.
XX
XX
AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
CC AB271479-AB271540 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening, and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing, treating, ameliorating or preventing
CC digestive disorders. Such conditions include disorders of the mouth,
CC oesophagus, stomach, small intestine, large intestine, liver, biliary
CC tract and pancreas, and include cancers of these organs and tissues. The
CC secreted proteins and their nucleic acids may also be used in the
CC treatment of immune disorders, inflammation, infection,
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
CC of the invention may be used for chromosome identification, chromosome
CC mapping, in gene therapy, for identifying individuals from minute
CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein-
CC encoding cDNA clone of the invention.
XX
SQ Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;
Query Match 98.9%; Score 1830.2; DB 25; Length 2077;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1832; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```


QY 1 GGTAGGCCGCGAGCTTAACTGCTGGAGCGGCTCCGTCGCCGCGCTGACAGCCGCTTA 60
Db 34 GGGTAGGCCGCGAGCTTAACTGCTGGAGCGGCTCCGTCGCCGCGCTGACAGCCGCTTA 93
QY 61 TCAGATTATCTTAACAAGAAAACCACTGGAAAAAATGAAATTCCTTATCTTCGAT 120
Db 94 TCAGATTATCTTAACAAGAAAACCACTGGAAAAAATGAAATTCCTTATCTTCGAT 153
QY 121 TTTTGGTGGTGTGACCTTTTATCCCTGTGCTGGGAAAGCTATATGCAAGATGGCA 180
Db 154 TTTTGGTGGTGTGACCTTTTATCCCTGTGCTGGGAAAGCTATATGCAAGATGGCA 213
QY 181 TCTTAAGAAGACTTTTGAAGAAATPAAAGAAATAGCCAGCTGTGGAGATGTTGCTA 240
Db 214 TCTTAAGAAGACTTTTGAAGAAATPAAAGAAATAGCCAGCTGTGGAGATGTTGCTA 273
QY 241 AAGCAATCATCACTAGCTGTTTATGTTAAAGCCCAAGAACGATCTTATGCGCATGG 300
Db 274 AAGCAATCATCACTAGCTGTTTATGTTAAAGCCCAAGAACGATCTTATGCGCATGG 333
QY 301 CACTTCGTGGTATGATCTGTGGAGCCAGACTGATGGCTCAAGAACTTGAAGAAAGCA 360
Db 334 CACTTCGTGGTATGATCTGTGGAGCCAGACTGATGGCTCAAGAACTTGAAGAAAGCA 393
QY 361 TCCAATTATATGACCAAAAACCTGACAGCAAGATGGCTGGAGAAAGTTCACTGAGCCAG 420
Db 394 TCCAATTATATGACCAAAAACCTGACAGCAAGATGGCTGGAGAAAGTTCACTGAGCCAG 453
QY 421 TGAAGAAATCCCACTGGAGAGAGGAGAAAGATCAGCTGTGATGCTGTGAGCCCAAGATTC 480
Db 454 TGAAGAAATCCCACTGGAGAGAGGAGAAAGATCAGCTGTGATGCTGTGAGCCCAAGATTC 513
QY 481 ATAAGATAGCCACTCTGGGCTTTGGCAGAGCATTGGGACTCCTCCAGAAAGGCTTAAAG 540
Db 514 ATAAATAGCCACTCTGGGCTTTGGCAGAGCATTGGGACTCCTCCAGAAAGGCTTAAAG 573
QY 541 CAGAAATCTGTGGTGTGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGCAAG 600
Db 574 CAGAAATCTGTGGTGTGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGCAAG 633
QY 601 GGAATATTGTTGTTTATTAACCACTTACATCACTCAAGAGAGGCTGCAATACGGA 660
Db 634 GGAATATTGTTGTTTATTAACCACTTACATCACTCAAGAGAGGCTGCAATACGGA 693
QY 661 CGCAGAGGAGCGGTGAGAGCTGCCAAGGTGGGGCTTTGGCATCTCATTTGATCCGAG 720
Db 694 CGCAGAGGAGCGGTGAGAGCTGCCAAGGTGGGGCTTTGGCATCTCATTTGATCCGAG 753
QY 721 CCTCCTTCTCATCTTACAGTCTCTCAACAGGTATTCAAGAAATACAGAGATGGCCCA 780
Db 754 CCTCCTTCTCATCTTACAGTCTCTCAACAGGTATTCAAGAAATACAGAGATGGCCCA 813
QY 781 AAATTCCAACAGCTGTATTTACGGGTGAAAGATGCAAAATGATGTCCAGAAATGGCTTCTC 840
Db 814 AGATTCCAACAGCTGTATTTACGGGTGAAAGATGCAAAATGATGTCCAGAAATGGCTTCTC 873
QY 841 ATGGATCAAAATTTGCTATTCAGCTTAAGATGGGGCAAAAGACTTACCAAGTACTGAT 900
Db 874 ATGGATCAAAATTTGCTATTCAGCTTAAGATGGGGCAAAAGACTTACCAAGTACTGAT 933
QY 901 CCTTCACAACCTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTGTACTGTCA 960
Db 934 CCTTCACAACCTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTGTACTGTCA 993
QY 961 GTGACATCTGGACAGCTGGGATGTTGGCAGGGGTGCCATGGATATGGCGGTGAGCCT 1020
Db 994 GTGACATCTGGACAGCTGGGATGTTGGCAGGGGTGCCATGGATATGGCGGTGAGCCT 1053
QY 1021 TTATATCATGGAGACACTCTCACTTATTAAGATCTTGGGCTGGTCCAAAGAGACTC 1080
Db 1054 TTATATCATGGAGACACTCTCACTTATTAAGATCTTGGGCTGGTCCAAAGAGACTC 1113
QY 1081 TGCGGCTGGTGTCTGCACTGCAAGAAACAAGGTGAGTGTGCTTCCAGTATATC 1140

Db 1114 TGCGGCTGGTGTCTGCACTGCAAGAAACAAGGTGAGTGTGCTTCCAGTATATC 1173
QY 1141 AGTAAACAAGTAAATATTTCCAACTACAGCTGTGGATGGAATGCAAGCAAGACT 1200
Db 1174 AGTAAACAAGTAAATATTTCCAACTACAGCTGTGGATGGAATGCAAGCAAGACT 1233
QY 1201 TCTTACCACACTGGCTCAATTCACCTGACAGTGAAGAGCCAGGACCATCATGAGAGG 1260
Db 1234 TCTTACCACACTGGCTCAATTCACCTGACAGTGAAGAGAGCCAGGACCATCATGAGAGG 1293
QY 1261 TTATGAGCTGTGCTGAGCCCTCAATATCACTCAAGTCTGAGCCATGAGAGAGGACAG 1320
Db 1294 TTATGAGCTGTGCTGAGCCCTCAATATCACTCAAGTCTGAGCCATGAGAGAGGACAG 1353
QY 1321 ACATCAACTTTTGGATCAAGCTGAGAGTGCCTGAGGCAAGCTTATGATGATATACA 1380
Db 1354 ACATCAACTTTTGGATCAAGCTGAGAGTGCCTGAGGCAAGCTTATGATGATATACA 1413
QY 1381 AGTATTTCTTCTTCATCACTCCACAGAGACACATGACTGTATGATGATCAAGAGCA 1440
Db 1414 AGTATTTCTTCTTCATCACTCCACAGAGACACATGACTGTATGATGATCAAGAGCA 1473
QY 1441 TGAATGTTGCTGTGCTGTGTTGGCTGTTTCTTATGTTGTTGCAATGAGAGAA 1500
Db 1474 TGAATGTTGCTGTGCTGTGTTGGCTGTTTCTTATGTTGTTGCAATGAGAGAA 1533
QY 1501 TGCTGCCATAGCTCTTAAACAGTAAAGAAAGAAAGTTTCAATGCTTGGCCAGGAATC 1560
Db 1534 TGCTGCCATAGCTCTTAAACAGTAAAGAAAGAAAGTTTCAATGCTTGGCCAGGAATC 1593
QY 1561 CTGGGCTGCAACTTGTGAAAAACTCTCTTCAACATTAACAATTCATCTTCA 1620
Db 1594 CTGGGCTGCAACTTGTGAAAAACTCTCTTCAACATTAACAATTCATCTTCA 1653
QY 1621 AAGCAACACTTATTCATGCTTCTGTATATCTTCTTGAATCTTCCAAATTCCT 1680
Db 1654 AAGCAACACTTATTCATGCTTCTGTATATCTTCTTGAATCTTCCAAATTCCT 1713
QY 1681 GATTCCTGAAAAGAAAGATTCATCCCTCCCTCCCAACATGATCAACATATGTA 1740
Db 1714 GATTCCTGAAAAGAAAGATTCATCCCTCCCTCCCAACATGATCAACATATGTA 1773
QY 1741 GGAATTCAGTGGGGGATTTCTTATATCACTCTTAAACAATGTTTCACTTTAA 1800
Db 1774 GGAATTCAGTGGGGGATTTCTTATATCACTCTTAAACAATGTTTCACTTTAA 1833
QY 1801 AGTAAACACTTATTAATTTTGGAGATCTTGA 1835
Db 1834 AGTAAACACTTATTAATTTTGGAGATCTTGA 1868
RESULT 12
AAK94491
ID AAK94491 standard; cDNA; 1860 BP.
XX AAK94491;
AC
XX
DT 06-NOV-2001 (first entry)
DE Human full-length cDNA, SEQ ID NO: 3328.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
OS Homo sapiens.
XX
PN EPI130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishit S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,
XX WPI; 2001-524255/58.
DR P-PSDB; AAM93559.
XX
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
XX
PS Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
XX molecules have been determined. Primers for synthesizing the full length
XX cDNA are useful for clarifying the function of the protein encoded by
XX the cDNA. The full length clones were obtained by construction of full
XX length enriched cDNA libraries that were synthesised by the oligo-capping
XX method. The primers enable the production of the full length cDNA easily
XX without any special methods. The present sequence is a full length
XX human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX
SQ Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;

Query Match 98.3%; Score 1820.4; DB 22; Length 1860;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1824; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGGTAGGCGCGAGCTTACCTGCTGGAGCGGCTTCGCGCGCGCTCAGAGCGCCCTA 60
DB 31 GGGTAGGCGCGAGCTTACCTGCTGGAGCGGCTTCGCGCGCGCTCAGAGCGCCCTA 90
QY 61 TCAGATATCTTAAACAGAAACCAACCTGGAAGAAAAATGAATTCCTTCTTCGAT 120
DB 91 TCAGATATCTTAAACAGAAACCAACCTGGAAGAAAAATGAATTCCTTCTTCGAT 150
QY 121 TTTTCGGTGTGTCACCTTTATCCCTGCTCTGGGAAGCTATATGCAAGATGCA 180
DB 151 TTTTCGGTGTGTCACCTTTATCCCTGCTCTGGGAAGCTATATGCAAGATGCA 210
QY 181 TCTCTAAGAGGACTTTGGAAGAAATAAGAAATAGCCAGCTGTGAGATGTTGCTA 240
DB 211 TCTCTAAGAGGACTTTGGAAGAAATAAGAAATAGCCAGCTGTGAGATGTTGCTA 270
QY 241 AAGCAATCATCACTAGCTGTTATGTGTAAGCCAGAAAGATCCTATAGCGATTGG 300
DB 271 AAGCAATCATCACTAGCTGTTATGTGTAAGCCAGAAAGATCCTATAGCGATTGG 330
QY 301 CACTTCGTGTGATACGTGTGAGCCAGAGCTGAGTGGCTCCAGAACTTAGAAAAAGCA 360
DB 331 CACTTCGTGTGATACGTGTGAGCCAGAGCTGAGTGGCTCCAGAACTTAGAAAAAGCA 390
QY 361 TCCAAATTTATGTACCAAACTCGAGCAAGATGGGCTGAGAAAGTTCACTGAGCCAG 420
DB 391 TCCAAATTTATGTACCAAACTCGAGCAAGATGGGCTGAGAAAGTTCACTGAGCCAG 450
QY 421 TGAAGATACCCCACTGGGAGAGGGAGAAAGATCACTGTATGCTGAGCCAGAAATTC 480
DB 451 TGAAGATACCCCACTGGGAGAGGGAGAAAGATCACTGTATGCTGAGCCAGAAATTC 510
QY 481 ATAAGATGCGATCCTGGGCTGTGGCAGCAGATTTGGGACTCTCCAGAAAGCAATTACAG 540
DB 511 ATAAGATGCGATCCTGGGCTGTGGCAGCAGATTTGGGACTCTCCAGAAAGCAATTACAG 570
QY 541 CAGAAGTTCTGTGGTGAACCTTTTCATGAACTGACAGAAAGGGCTCAGAAAGCAAG 600

DB 571 CAGAAGTTCTGTGGTGAACCTTTTCATGAACTGACAGAAAGGGCTCAGAAAGCAAG 630
QY 601 GGAAGATTTGTTTATTAACCAACTTACATCACTCAAGAGCGGTCAATACCGAA 660
DB 631 GGAAGATTTGTTTATTAACCAACTTACATCACTCAAGAGCGGTCAATACCGAA 690
QY 661 CGCAGGGGGGGTGAAGCTGCCAAGGTGGGGGCTTTGGCATCTTCATTCGATCCGTGG 720
DB 691 CGCAGGGGGGGTGAAGCTGCCAAGGTGGGGGCTTTGGCATCTTCATTCGATCCGTGG 750
QY 721 CTTCTTCTCATCTACAGTCTCTCAACAGATATTCAGAAATACAGAGATGGCTGCCA 780
DB 751 CTTCTTCTCATCTACAGTCTCTCAACAGATATTCAGAAATACAGAGATGGCTGCCA 810
QY 781 AAATTCACAGGCTGTATTAAGGTGGAATGAGAAATGATGTCAAGATGGCTCTC 840
DB 811 AGATTCACAGGCTGTATTAAGGTGGAATGAGAAATGATGTCAAGATGGCTCTC 870
QY 841 ATGGATCAAAATTTGTCATTCAGCTAAAGTGGGGGCAAGACCTACAGATACGATT 900
DB 871 ATGGATCAAAATTTGTCATTCAGCTAAAGTGGGGGCAAGACCTACAGATACGATT 930
QY 901 CTTCAACTGTGACAGAGATCACTGGAGCAAAATTCAGAAACAGTTGTACTGTCA 960
DB 931 CTTCAACTGTGACAGAGATCACTGGAGCAAAATTCAGAAACAGTTGTACTGTCA 990
QY 961 GTGACATCTGACAGCTGGGATGTTGGGAGGTGTCAGAGATGGCGGTGAGCCT 1020
DB 991 GTGACATCTGACAGCTGGGATGTTGGGAGGTGTCAGAGATGGCGGTGAGCCT 1050
QY 1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTC 1080
DB 1051 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTC 1110
QY 1081 TGGGCTGTGCTCTGAGCTGACAGAGAACAAAGTGGAGTTGGCTTCCAGTATATC 1140
DB 1111 TGGGCTGTGCTCTGAGCTGACAGAGAACAAAGTGGAGTTGGCTTCCAGTATATC 1170
QY 1141 AGTTACCAAGGTAAATTTTCAACACAGCTGTGGATGAGATCGACGAGGAACT 1200
DB 1171 AGTTACCAAGGTAAATTTTCAACACAGCTGTGGATGAGATCGACGAGGAACT 1230
QY 1201 TCTTACCACTGGGCTCAATTCACCTGAGGAAAGGCAAGGCGCATATGGAGAGG 1260
DB 1231 TCTTACCACTGGGCTCAATTCACCTGAGGAAAGGCAAGGCGCATATGGAGAGG 1290
QY 1261 TTATGAGCTGTGACAGCCCTCAATATCACTCAAGTCTGAGCCATGAGAAAGGACAG 1320
DB 1291 TTATGAGCTGTGACAGCCCTCAATATCACTCAAGTCTGAGCCATGAGAAAGGACAG 1350
QY 1321 ACATCACTTTTGGATTCAGAGTGGAGTGGCTGAGGCACTTACTGATGACTTATACA 1380
DB 1351 ACATCACTTTTGGATTCAGAGTGGAGTGGCTGAGGCACTTACTGATGACTTATACA 1410
QY 1381 AGTATTTCTTCTTCATCACTCCACAGGAGACACATGATGATGATCCAAAGAGA 1440
DB 1411 AGTATTTCTTCTTCATCACTCCACAGGAGACACATGATGATGATCCAAAGAGA 1470
QY 1441 TGAATGTTGCTGTGCTGTTTGGGCTGTTTCTTATGTTGTTGACAGATGGAGAAA 1500
DB 1471 TGAATGTTGCTGTGCTGTTTGGGCTGTTTCTTATGTTGTTGACAGATGGAGAAA 1530
QY 1501 TGTGCTAGGCTCTTAAAGACAGTAAAGAAAGGTTTTCATGCTTCTGCGCAGAAATC 1560
DB 1531 TGTGCTAGGCTCTTAAAGACAGTAAAGAAAGGTTTTCATGCTTCTGCGCAGAAATC 1590
QY 1561 CTGGGCTGCAACTTTGAGAAAGCTCTTCAACATTAATTTATCATCACTCATCTTCA 1620
DB 1591 CTGGGCTGCAACTTTGAGAAAGCTCTTCAACATTAATTTATCATCACTCATCTTCA 1650
QY 1621 AAGCAAACTTATTTATGCTTGTATATATCTTGTATATCTTGTATATCTTCAATCTCT 1680
DB 1651 AAGCAAACTTATTTATGCTTGTATATATCTTGTATATCTTGTATATCTTCAATCTCT 1710

Oy	1661	GATTCGTAGAAAAAGGAATCATCTCCCTCCCTCCACACATAGATGAATCAATATGGTA	1776
Db	1711	GATTTTGAHAAAAAGGAAATCATCTCCCTCCCTCCACACATAGATCAATATGGTA	1776
Oy	1741	GGGATTACAGTGGGGGCACTTTCTTTATATACCTCTTAAAAAATTTGTTCCACTTTAA	1800
Db	1771	GGGATTACAGTGGGGGCACTTTCTTTATATACCTCTTAAAAAATTTGTTCCACTTTAA	1833
Oy	1801	AGTAACACTTATATTAATTTTGGGAAGATC	1830
Db	1831	AGTAACACTTATATTAATTTTGGGAAGATC	1860
RESULT 13			
ID	AAAA0493	standard; cDNA; 1767 BP.	
XX	AAAA0493;		
XX	16-NOV-2000	(first entry)	
XX	Human fetal kidney cDNA fragment AM282_11.		
XX	Secreted protein; cytosolic; immunostimulatory; antimicrobial;		
XX	antiviral; immunosuppressive; antiinflammatory; vulnerrary; cytokine;		
XX	cell proliferation; differentiation; regulator; treatment; tumor;		
XX	autoimmune disease; inflammatory disorder; wound; microbial infection;		
XX	vital disease; graft versus host reaction suppression; ss.		
XX	Homo sapiens.		
XX	MO200037630-A1.		
XX	29-JUN-2000.		
XX	22-DEC-1999;	99WO-US31005.	
XX	23-DEC-1998;	98US-0220876.	
XX	(GENY) GENETICS INST INC.		
XX	Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;		
XX	Merberg D, Treacy M, Bowman MR;		
XX	WPI; 2000-442661/38.		
XX	P-PSDB; AAB10229.		
XX	Secreted human proteins AS296-1i and AS34-1i, useful for treating		
XX	tumors, autoimmune diseases, inflammatory disorders, wounds, microbial		
XX	infections and viral diseases -		
XX	Disclosure; Page 198; 293pp; English.		
XX	This invention describes novel secreted human proteins (I) which have		
XX	cytostatic, immunostimulatory, antimicrobial, antiviral,		
XX	immunosuppressive, antiinflammatory and vulnerary activity and which act		
XX	as cytokine, cell proliferation or differentiation regulators. (I)		
XX	is useful for treating tumors, autoimmune diseases, inflammatory		
XX	disorders, wounds, microbial infections and viral diseases. (I) is also		
XX	useful for suppressing graft versus host reaction. AAA40490-A40580		
XX	represent cDNA fragments that encode the secreted proteins		
XX	AA10226-110288 described in the method of the invention.		
XX	Sequence 1767 BP; 512 A; 384 C; 418 G; 453 T; 0 other;		
XX	Query Match	94.6%; Score 1751.2; DB 21; Length 1767;	
XX	Best Local Similarity	99.8%; Pred. No. 0;	
XX	Matches 1753; Conservative	0; Mismatches 3; Indels 0; Gaps 0;	
Oy	96	AAATGAATTCCTATCTCGCAATTTTGGTGTGATGTCACCTTTATCCCTGTGCT	155
Db	3	AAGATGAATTCCTATCTCGCAATTTTGGTGTGATGTCACCTTTTATCCCTGTGCT	62

QY	156	GGGAAAGCTATATGCAAGATATGGCATCTCTTAAGAGGACTTTTGAAGAAATTAAGAAGAA	215
Db	63	GGGAAAGCTATATGCAAGATATGGCATCTCTTAAGAGACTTTTGAAGAAATTAAGAAGAA	122
QY	216	ATAGCAGCTGTGAGATATGTTGCTTAAGCAATCATCAACCTAGCTGTTATGGTAAAGCC	275
Db	123	ATAGCAGCTGTGAGATATGTTGCTTAAGCAATCATCAACCTAGCTGTTATGGTAAAGCC	182
QY	276	CAGAAAGATCCTATAGACCATGGCACTTCGTGGTATATACGTGGAGCCCACTGAGT	335
Db	183	CAGAAAGATCCTATAGACCATGGCACTTCGTGGTATATACGTGGAGCCCACTGAGT	242
QY	336	GGCTCAAGAACTTAGAAAAAGCCATCCAAATTTGTATACAAAACCTGCAGCAAGATGGG	395
Db	243	GGCTCAAGAACTTAGAAAAAGCCATCCAAATTTGTATACAAAACCTGCAGCAAGATGGG	302
QY	396	CTGAGAAAGTTCACTCTGAGACCTAGTGAATATCCCACTGGAGAGGGAGAAATCA	455
Db	303	CTGAGAAAGTTCACTCTGAGACCTAGTGAATATCCCACTGGAGAGGGAGAAATCA	362
QY	456	GCCTGATCTGTGAGACCAAGAAATTCATATGATATACCATCCGGGCTCTGGCAGACATTT	515
Db	363	GCCTGATCTGTGAGACCAAGAAATTCATATGATATACCATCCGGGCTCTGGCAGACATTT	422
QY	516	GGGACTCTTCAGAAAGGCATTTACAGCAGAAATTTGTGGTGAACCTTTTGATGAATG	575
Db	423	GGGACTCTTCAGAAAGGCATTTACAGCAGAAATTTGTGGTGAACCTTTTGATGAATG	482
QY	576	CAGAGAAAGGCTCAGAAAGCAAGAGGAAATTTGTTTATACCAACTTATCATCAAC	635
Db	483	CAGAGAAAGGCTCAGAAAGCAAGAGGAAATTTGTTTATACCAACTTATCATCAAC	542
QY	636	TACTCAAGACCGTGCATATACGAAACGCAAGGGGCGGTGGAAGCTGCCAAGTGGGGCT	695
Db	543	TACTCAAGACCGTGCATATACGAAACGCAAGGGGCGGTGGAAGCTGCCAAGTGGGGCT	602
QY	696	TTGGCATCTCTATTCGATCCGTGGCTCTTTCTTCACTTACAGTCTCAACAAGTATT	755
Db	603	TTGGCATCTCTATTCGATCCGTGGCTCTTTCTTCACTTACAGTCTCAACAAGTATT	662
QY	756	CAGGAATACAGAGATGGGGTGGCCAAAATTTCCACACCTGTATTAAGGTGAAGATGCA	815
Db	663	CAGGAATACAGAGATGGGGTGGCCAAAATTTCCACACCTGTATTAAGGTGAAGATGCA	722
QY	816	GAATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGTCAATCAGCTAAAGATGGG	875
Db	723	GAATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGTCAATCAGCTAAAGATGGG	782
QY	876	GCAAAACCTACCCAGATATAGATTTCCCTTCAACCTGTACAGAGATCACTGGAGCAAA	935
Db	783	GCAAAACCTACCCAGATATAGATTTCCCTTCAACCTGTACAGAGATCACTGGAGCAAA	842
QY	936	TATTCAGAAACAGTTGTACTGGTCAGTGAACATCTGCAACAGCTGGATGTGGGACGGT	995
Db	843	TATTCAGAAACAGTTGTACTGGTCAGTGAACATCTGCAACAGCTGGATGTGGGACGGT	902
QY	996	GCCATGATGATGGCGGTGAGACCTTTATATCATGGGAACACTCTCACTTATTAAGAT	1055
Db	903	GCCATGATGATGGCGGTGAGACCTTTATATCATGGGAACACTCTCACTTATTAAGAT	962
QY	1056	CTTGGGCTGTGCAAGAGGACTCTGCGGTGGTGTCTGGACTGCAAGAAACAAGGT	1111
Db	963	CTTGGGCTGTGCAAGAGGACTCTGCGGTGGTGTCTGGACTGCAAGAAACAAGGT	102
QY	1116	GGAGTTGGTGCCTTCAGATATTTATCATGTTACAAAGGTAAATTTTCCACTACAGTCTG	117
Db	1023	GGAGTTGGTGCCTTCAGATATTTATCATGTTACAAAGGTAAATTTTCCACTACAGTCTG	108
QY	1176	GTGATGATCTGCAAGCAAGAAACCTTCTTACCCACTGGGCTGCAATTCACCTGGCACTGAA	1231
Db	1083	GTGATGATCTGCAAGCAAGAAACCTTCTTACCCACTGGGCTGCAATTCACCTGGCACTGAA	1144

Db 423 GCACGACGATTGGGACCTCTCCAGAGGCAATTACAGAGAAGTTCTGTGATGACCTCTT 482
 Qy 565 TCGATGAACTGCAAGAGAGGCTTCAGAGCAAGGGAGAAATGTTTATTAACACAC 624
 Db 483 TCGATGAACTGCAAGAGAGGCTTCAGAGCAAGGGAGAAATGTTTATTAACACAC 542
 Qy 625 CTTCATCACTACTCAAGAGAGGCTTCAGAGCAAGGGAGAAATGTTTATTAACACAC 684
 Db 543 CTTCATCACTACTCAAGAGAGGCTTCAGAGCAAGGGAGAAATGTTTATTAACACAC 602
 Qy 685 AGGTGGGGGCTTTGGGCACTCTCATTCGATCCGCTCTCTTCTCATCTACAGTCTCT 744
 Db 603 AGGTGGGGGCTTTGGGCACTCTCATTCGATCCGCTCTCTTCTCATCTACAGTCTCT 662
 Qy 745 AACAGGATTCAGGAAATACAGAGATGGGCGGCCCAAAATTCACACAGCTGATTAAGG 804
 Db 663 AACAGGATTCAGGAAATACAGAGATGGGCGGCCCAAAATTCACACAGCTGATTAAGG 722
 Qy 805 TGAAGATGCAAGAAATGATGTCAGAAATGAGCTTCATGAGGATCAAAATGTCATTCAGC 864
 Db 723 TGAAGATGCAAGAAATGATGTCAGAAATGAGCTTCATGAGGATCAAAATGTCATTCAGC 782
 Qy 865 TAAAGATGGGGCAAAAGACTTACCCAGATTAATTCCTTCAACCTGTAGCAGAGATCA 924
 Db 783 TAAAGATGGGGCAAAAGACTTACCCAGATTAATTCCTTCAACCTGTAGCAGAGATCA 842
 Qy 925 CTGGAGCAAAATATCCAGAAACAGGTTTGTACTGTGTAGTGAGCATGTCGACAGCTGGGATG 984
 Db 843 CTGGAGCAAAATATCCAGAAACAGGTTTGTACTGTGTAGTGAGCATGTCGACAGCTGGGATG 902
 Qy 985 TTGGGCAAGGTCCTCATGATGATGAGCTGAGGCTTATATATCATGAGGAGCACTCTCAC 1044
 Db 903 TTGGGCAAGGTCCTCATGATGATGAGCTGAGGCTTATATATCATGAGGAGCACTCTCAC 962
 Qy 1045 TTATTTAAAGATCTTGGGCTGCTGCTCAAAAGAGACTCTGCGGCTGTGCTGTGAGCTGAG 1104
 Db 963 TTATTTAAAGATCTTGGGCTGCTGCTCAAAAGAGACTCTGCGGCTGTGCTGTGAGCTGAG 1022
 Qy 1105 AAGAACAGGTCGATGTTGCTGCTCCAGTATTAATCACTGTAACAAAGGTAATATTTTCCA 1164
 Db 1023 AAGAACAGGTCGATGTTGCTGCTCCAGTATTAATCACTGTAACAAAGGTAATATTTTCCA 1082
 Qy 1165 ACTACAGCTGCGGTCGATGATGTCGAGCAGAGAACTTCTTATCCCACTGGGCTGCATTTCA 1224
 Db 1083 ACTACAGCTGCGGTCGATGATGTCGAGCAGAGAACTTCTTATCCCACTGGGCTGCATTTCA 1142
 Qy 1225 CTGGCAGTGAAGAGGCCAGGCGCATCATGAGAGAGGTTATGAGCTGTGACGCCCTCA 1284
 Db 1143 CTGGCAGTGAAGAGGCCAGGCGCATCATGAGAGAGGTTATGAGCTGTGACGCCCTCA 1202
 Qy 1285 ATATCACTCAGGTCCTGAGCCATGAGAAAGGAGCAAGCACTTTTGGATCCAAAGCTG 1344
 Db 1203 ATATCACTCAGGTCCTGAGCCATGAGAAAGGAGCAAGCACTTTTGGATCCAAAGCTG 1262
 Qy 1345 GAGTGCCTGAGCCAGTCTACTTATGATGATTAACAAGATTTCTTCTTCCATCACTGCC 1404
 Db 1263 GAGTGCCTGAGCCAGTCTACTTATGATGATTAACAAGATTTCTTCTTCCATCACTGCC 1322
 Qy 1405 ACGGAGACCACTGATCTGTATGATCCAAAGCAGATGAATGTTGCTGTGTTTGGG 1464
 Db 1323 ACGGAGACCACTGATCTGTATGATCCAAAGCAGATGAATGTTGCTGTGTTTGGG 1382
 Qy 1465 CTGTGTTTCTTATGTTGTGAGCATGAGAAAGAAATGCTGCTAGGCTCTTGAAGAAAGCT 1524
 Db 1383 CTGTGTTTCTTATGTTGTGAGCATGAGAAAGAAATGCTGCTAGGCTCTTGAAGAAAGCT 1442
 Qy 1525 AAGAAAGAAAGCTTTTCAATGCTTCTGGGCAAGATCTGGGCTGCAACTTTGAGAAACT 1584
 Db 1443 AAGAAAGAAAGCTTTTCAATGCTTCTGGGCAAGATCTGGGCTGCAACTTTGAGAAACT 1502
 Qy 1585 CCTCTTCAATTAATTTCAATTCATCTTCAAGCACTGATTTTCAATGCTTT 1644
 Db 1503 CCTCTTCAATTAATTTCAATTCATCTTCAAGCACTGATTTTCAATGCTTT 1562

Qy 1645 CTGTATATATCTTCTTCTGATACCTTTCCAAATTCCTGATTTCTAGAAAAAGAAATCATTTCT 1704
 Db 1563 CTGTATATATCTTCTTCTGATACCTTTCCAAATTCCTGATTTCTAGAAAAAGAAATCATTTCT 1622
 Qy 1705 CCCCTCCCTCCACACATAGAAATCAACATATGATGATGAGATTAACATGAGGAGCAATTTCTT 1764
 Db 1623 CCCCTCCCTCCACACATAGAAATCAACATATGATGATGAGATTAACATGAGGAGCAATTTCTT 1682
 Qy 1765 TATATCACTCTTAAACATTTGTTTCCATTTTAAAGTAAACCTTAATTAATTTTGG 1824
 Db 1683 TATATCACTCTTAAACATTTGTTTCCATTTTAAAGTAAACCTTAATTAATTTTGG 1742
 Qy 1825 AAGATCTGTGAAAAAATTTTAAAAA 1851
 Db 1743 AAGATCTGTGAAAAAATTTTAAAAA 1769
 RESULT 15
 AAH9703
 ID AAH9703 standard; cDNA; 1895 BP.
 XX
 AC AAH9703;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human protein encoding cDNA sequence SEQ ID NO:538.
 XX
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiant; central nervous system; vitruide;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KW antiaggregant; haemostatic; vulnery; antilucer; osteopathic; eczema;
 KW dermatological; antiallergic; antidiabetic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiParkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antinaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; hematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PE 22-DEC-2000; 2000WO-US35017.
 XX
 PR 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 XX
 PA (HVSF-) HVSFQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 PI
 XX
 DR WPI; 2001-457603/49.
 DR
 DR P-PSDB; AAM25762.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX
 XX
 PS Claim 1; Page 591; 1217p; English.
 CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAH25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;

CC cardiovascular; antihaemic; antiagregant; haemostatic; vulnery;
CC antiulcer; osteoplastic; dermatological; antiallergic; antiaethmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
encoding them can be used in gene therapy, antisense therapy and vaccine
production. The proteins and polynucleotides are useful for screening for
agonists or antagonists of a protein and for the treatment and diagnosis
of disorders associated with the activity of a protein e.g. inflammation,
rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
infections, autoimmunity, genetic diseases, haematopoietic disorders,
anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
osteoporosis, severe combined immunodeficiency, eczema, allergic
rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.

XX Sequence 1895 BP, 530 A; 439 C; 450 G; 476 T; 0 other;

Query Match 87.4%; Score 1617.2; DB 22; Length 1895;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1811; Conservative 0; Mismatches 33; Indels 24; Gaps 15;

QY 1 GGTAGGCGCGAGCTTATGTCCTGGAGCGCCTTCCTGCGCCGCTGAGAGCGCCCTTA 60
DB 28 GGTAGGCGCGAGCTTATGTCCTGGAGCGCCTTCCTGCGCCGCTGAGAGCGCCCTTA 87
QY 61 TCAGATTATCTTAACAAGAAACCACTGGAAGAAAAAATGAATTCCTTCTTCGAT 120
DB 88 TCAGATTATCTTAACAAGAAACCACTGGAAGAAAAAATGAATTCCTTCTTCGAT 147
QY 121 TTTTCGGTGGTGTACCTTTATCCCTGTGCTGGAAGAGTATATGCAAGATGCA 180
DB 148 TTTTCGGTGGTGTACCTTTATCCCTGTGCTGGAAGAGTATATGCAAGATGCA 207
QY 181 TCTCTAAGAGAGCTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTTGCTA 240
DB 208 TCTCTAAGAGAGCTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTTGCTA 267
QY 241 AAGGAATCATCAACTGAGCTTTTATGTTAAGCCCAAGACAGATCCATATAGGATGG 300
DB 268 AAGGAATCATCAACTGAGCTTTTATGTTAAGCCCAAGACAGATCCATATAGGATGG 327
QY 301 CACTTCTGTGATGACTGTGAGCCAGACTGAGTGGCTCCAAGAACTTAGAAAAAGCA 360
DB 328 CACTTCTGTGATGACTGTGAGCCAGACTGAGTGGCTCCAAGAACTTAGAAAAAGCA 387
QY 361 TCCAAATTATGTACCAAAACCTGACAGCAAGATGGGCTGGAAGAACTTCACTGAGCCAG 420
DB 388 TCCAAATTATGTACCAAAACCTGACAGCAAGATGGGCTGGAAGAACTTCACTGAGCCAG 447
QY 421 TGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTATGCTGAGCCAGAAATTC 480
DB 448 TGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTATGCTGAGCCAGAAATTC 507
QY 481 ATAGATATGCACTCTGGGCTTGTGGACAGACTTGGGACTCTCCAGAAAGCAATTACAG 540
DB 508 ATAGATATGCACTCTGGGCTTGTGGACAGACTTGGGACTCTCCAGAAAGCAATTACAG 567
QY 541 CAGAAATTCCTGGTGTGACTCTTTCGATGAACTGCAAGAGAGGCTCAGAAAGCAGAG 600
DB 568 CAGAAATTCCTGGTGTGACTCTTTCGATGAACTGCAAGAGAGGCTCAGAAAGCAGAG 627
QY 601 GGAAGATTTGTTTATTAACAACCTTACATCACTCAAGAGCGGTGCAATACCGAA 660
DB 628 GGAAGATTTGTTTATTAACAACCTTACATCACTCAAGAGCGGTGCAATACCGAA 687
QY 661 CGCAGGGGGGGTGGAGCTGCCAAGGTGGGGGCTTGGCATCTCTCAATTCGATCCGGG 720
DB 688 CGCAGGGGGGGTGGAGCTGCCAAGGTGGGGGCTTGGCATCTCTCAATTCGATCCGGG 747
QY 721 CCTCTCTCCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAGATGGCGTCCCA 780

DB 748 CCTCTCTCCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAGATGGCGTCCCA 807
QY 781 AATATTCACAGCGCTGATTAACGGTGGAGATGAGCAAGATGATCAAGATGGCTCTC 840
DB 808 AGATTCCACAGCGCTGATTAACGGTGGAGATGAGCAAGATGATCAAGATGGCTCTC 867
QY 841 ATGGGATCAAAATTTGTCATTCACTTAAGATGGGGGCAAGACCTTACCAAGATCTGAT 900
DB 868 ATGGGATCAAAATTTGTCATTCACTTAAGATGGGGGCAAGACCTTACCAAGATCTGAT 927
QY 901 CTTTCAACACTGTATGCAAGATCACTGTGGCAAAATTTCCAGAAACAGTTGTACTGTCA 960
DB 928 CTTTCAACACTGTATGCAAGATCACTGTGGCAAAATTTCCAGAAACAGTTGTACTGTCA 987
QY 961 GTGACATCTGGACAGTGGGAGTGTGGAGGGGTGCATGAGATGGGGGTGGAGCT 1020
DB 988 GTGACATCTGGACAGTGGGAGTGTGGAGGGGTGCATGAGATGGGGGTGGAGCT 1047
QY 1021 TTATATCATGGGAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGCACTC 1080
DB 1048 TTATATCATGGGAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGCACTC 1107
QY 1081 TGGCGTGTGCTCTGCACTGCAAGAACCAAGGTGAGTGTGCTTCCAGTATATC 1140
DB 1108 TGGCGTGTGCTCTGCACTGCAAGAACCAAGGTGAGTGTGCTTCCAGTATATC 1167
QY 1141 AGTTACCAAGGTAAATTTCCAACTACAGCTGTGATGAGATCTGACGCAAGAACT 1200
DB 1168 AGTTACCAAGGTAAATTTCCAACTACAGCTGTGATGAGATCTGACGCAAGAACT 1227
QY 1201 TCTTACCACCTGGGCTCAATTCAGTGGCAAGAAAGCCAGGGCCATCATGAGAGAG 1260
DB 1228 TCTTACCACCTGGGCTCAATTCAGTGGCAAGAAAGCCAGGGCCATCATGAGAGAG 1287
QY 1261 TTATGACCTGTGCTGACGCTCAATATCACTCAGTCTGAGCCATGAGAGAGAG 1320
DB 1288 TTATGACCTGTGCTGACGCTCAATATCACTCAGTCTGAGCCATGAGAGAGAG 1347
QY 1321 ACATCAACTTTGGATCCAGCTGAGATGCTGAGACAGTCTGATGATGATCA 1380
DB 1348 ACATCAACTTTGGATCCAGCTGAGATGCTGAGACAGTCTGATGATGATCA 1407
QY 1381 AGTATTTCTTCTTCATCACTCCAGGAGACACATGACTGTC---ATGATCCAAAGC 1437
DB 1408 AGTATTTCTTCTTCATCACTCCAGGAGACACATGACTGTC---ATGATCCAAAGC 1467
QY 1438 AGATGAATG-TTGGCTGTGCTG-TTGGGCTGTGTTCTTATG-TGTTGACAGATGG 1494
DB 1468 AGATGAATG-TTGGCTGTGCTG-TTGGGCTGTGTTCTTATG-TGTTGACAGATGG 1527
QY 1495 AAGAAATGCTGCTGAGTCTTGAAGAACATGAAGAAAGAA-CGTTTTCATGCTTGTGCC 1553
DB 1528 AAGAAATGCTGCTGAGTCTTGAAGAACATGAAGAAAGAA-CGTTTTCATGCTTGTGCC 1587
QY 1554 --AGAAATGCTGCTGCAACTTT-GAAAAACCTCTTCACATCAATTT--CATCC 1608
DB 1588 CAGGAATTCCTGGGCTGTGCACTTGGGAAAAACCTCTTCACATCAATTT--CATCC 1647
QY 1609 AATTCATCTTCAAGCACACTCT-ATTCAATGCTTCT-GTATTAATCTTCTT-GATA 1665
DB 1648 AATTCATCTTCAAGCACACTCTTAATTCATGCTTCTGTTATATCTTCTTGGATA 1707
QY 1666 CTTTCCAAATTTCTG--ATTCTGAAGAAAGAAATCAATTTCTCCCTCC--CTCCACC 1719
DB 1708 CTTTCCAAATTTCTGATTTCTGAAGAAAGGGAATCAATTTCTCCCTCCCTCCACC 1767
QY 1720 ACATAGATCAACATATGATGAGGATTAACAGTGGGGGCA-TTCTTATATACCTCTTA 1778
DB 1768 ACATAGATCAACATATGATGAGGATTAACAGTGGGGGCA-TTCTTATATACCTCTTA 1827
QY 1779 AAAACATGTTTCCACTTTAAA--GTAAACATTAATTAATTTTGAAGATCTGTAA 1836
DB 1828 AAAACATGTTTCCACTTTAAAAGTTAAACATTAATTAATTTTGAAGATTAATCTGAA 1887

Oy 1837 AAAAAAA 1844
| | | | |
Db 1886 AAAAAAA 1895

Search completed: December 22, 2003, 15:09:56
Job time : 455.127 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:22:34 ; Search time 3623 Seconds
(without alignments)
12417.225 Million cell updates/sec

Title: US-09-745-763-35
Perfect score: 1851
Sequence: 1 GGCTAGGCGCCGAGCTTAGT.....CTGAAAAAAAAAAAAAAAAA 1851

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estdb: *
2: em_esthum: *
3: em_estm: *
4: em_estmu: *
5: em_estcov: *
6: em_estpl: *
7: em_estro: *
8: em_hlc: *
9: gb_est1: *
10: gb_est2: *
11: gb_hlc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vrt: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rtd: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gss1: *
29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1164.8	62.9	1739	AK032972 Mus muscu
2	1164	62.9	1868	BC010977 Mus muscu
3	1162.2	62.8	1805	AK075686 Mus muscu
4	993.8	53.7	1024	BC012019 Homo sapi

C	5	992.8	53.6	996	13	BX355939	BX355939
	6	985.8	53.3	1050	13	BX439467	BX439467
	7	940.8	50.8	1038	13	BX355940	BX355940
C	8	925.6	50.0	1020	13	BX439466	BX439466
	9	894	48.3	1073	13	BX360506	BX360506
	10	872.4	47.1	1126	13	BX360507	BX360507
	11	847.8	45.7	909	13	BQ878966	BQ878966
C	12	845.2	45.8	1003	13	BX416895	BX416895
	13	840.8	45.4	1012	13	BX416896	BX416896
	14	838.6	45.3	919	13	BX335995	BX335995
C	15	831.6	44.9	929	13	BX335994	BX335994
	16	831	44.9	998	13	BX460463	BX460463
	17	772.6	41.7	945	12	B1909780	B1909780
	18	761.6	41.1	801	13	B1754143	603025609
	19	753.8	40.7	1081	13	BQ072892	AGENCOURT
	20	748.6	40.4	866	13	B1146905	AGENCOURT
	21	745.4	40.3	922	12	BG761741	602717936
	22	740	40.0	902	10	BG751497	602730076
	23	739.8	40.0	783	12	B1754468	603022607
	24	724.2	39.1	879	13	BQ218838	AGENCOURT
C	25	723.2	39.1	752	14	CD370462	CD370462
	26	719.8	38.9	832	12	B1911772	UI-H-FT1-
	27	719.2	38.9	739	14	CB958693	AGENCOURT
	28	716.4	38.7	786	10	BG430966	BG430966
	29	709	38.3	944	12	B1819273	603037713
	30	691.4	37.4	774	14	CB308109	AGENCOURT
	31	689.8	37.3	1201	13	BX379674	BX379674
	32	688.8	37.2	806	12	B1764052	603043258
C	33	686.8	37.1	703	13	B0625199	UI-H-FG1-
	34	684	37.0	805	10	BG563740	602584524
	35	683.6	36.9	813	14	CD519131	AGENCOURT
C	36	681.8	36.8	708	14	CD366462	UI-H-FT1-
	37	676.4	36.5	691	14	CA436850	UI-H-FT1-
C	38	675.8	36.5	701	12	B1771109	603059666
	39	675.8	36.5	745	12	B1462884	603203005
	40	671.6	36.3	729	12	B1460469	603200966
	41	665.2	35.9	889	10	BE906771	601498329
	42	663.8	35.9	937	14	CD515740	AGENCOURT
	43	662.2	35.2	727	12	BG765433	602738911
	44	638.2	34.5	807	10	BG190926	RST10008
C	45	636.8	34.4	653	10	BE218907	hV46D06.X

ALIGNMENTS

RESULT 1
AK032972
LOCUS
DEFINITION Mus musculus 12 days embryo male wolffian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:6720483C06
ACCESSION AK032972
VERSION AK032972.1 GI:26328732
KEYWORDS product:Plasma glutamate carboxypeptidase, full insert sequence.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
2 99279253 MEDLINE
10349636 PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
JOURNAL
MEDLINE
PUBMED
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
11042159

REFERENCE AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, K., Fujimake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nakado, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Flacher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzei, J., Nombela, P., Nordone, P., Rising, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlschki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
JOURNAL REFERENCE AUTHORS	6 (bases 1 to 1739) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sueno-cho, Tsukuba-shi, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ location/Qualifiers
FEATURES	

source	1. 1739 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:6720483C06" /db_xref="taxon:10090" /clone="6720483C06" /sex="male" /tissue_type="olfactory duct; includes surrounding region" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="12 days embryo" 81. 1493 /note="unnamed protein product; plasma glutamate carboxypeptidase (MCD) MGI:1889205, GB NM_018755, evidence: BLASTN, 99%, match=543) putative"
CDS	/codon_start=1 /protein_id="BAC28105.1" /db_xref="GI:26328733" /translation="MRSIFPLFVHLALGSGKAVPKNGVORTPREIKELIANVEDY AKALINAVGKXONSVERLGLVDVVGRLSGSKRLKALIMONLOODGLENYH LEQVRIHMERGESAVMEPRIRHRAILGLSSISGTPREGITAEVLVASFPELQAR ASEARGLIVNOPYTEYEVQVROGAVEAKVGLVSLIOSVASFISPHGTIO KYDGVKPIPLACITVEDAEEMSRMSRGNKIVIHLEMGAKTVPDTSFNTVAITGS MYBEVVLVGHLDSDMDVGQALDDGGAFISWALSLVQLGLRPRTRILVLTAE EQQGIGASQYVEHLKANIISKYSLVMEADSGPLPTGLQPTGSKARAKIMKVNLLQF LNTKVSNEGCTIDNPVIOAGVPGASLRDILKYFFHHSHDITVMDPKQNVAA AYAAVAVVYVADMDENLPRS" 1722. 1737 /note="putative" 1739 /note="putative"
polya_signal	
polya_site	
BASE COUNT	482 a 369 c 423 g 465 t
ORIGIN	
Query Match	62.9%; Score 1164.8; DB 11; Length 1739;
Best Local Similarity	83.8%; Pred. No. 3.1e-211;
Matches 1329; Conservative 0; Mismatches 254; Indels 3; Gaps 1;	
131 TGTTCACCTTTTATTCCTGTGCTGGGAAAGCTATATGCAAGATGCACTTTAAGAG	190
107 TGTTCACCTTTTATTCCTGTGCTGGGAAAGCTATATGCAAGATGCTTTCTCAGCG	166
191 GACTTTGAAGAATAAAGAAGAAATAGCAGCTGGAGATGCTTCTAAGCAATCAT	250
167 AACCTTTCGAAATTAAGAAAGAAATAGCAGCTTGAAGATGCTTCTAAGCAATCAT	226
251 CAACCTAGCTGTTTATGTTAAGCCAGAAACAATCCTATAGAGATGCACTTCTGCT	310
227 CAACCTAGCTGTTTATGTTAAGCCAGAAACAATCCTATAGAGATGCACTTCTGCT	286
311 TGATACCTTTGAGACCCAGACTGAGTGGCTCCAGAACTTGAAGAAAGCAATCAATAT	370
287 TGATACCTTTGAGACCCAGACTGAGTGGCTCCAGAACTTGAAGAAAGCAATCAATAT	346
371 GTACCAAAACCTGAGCAAGATGAGGCTGAGAAAGTTCACTGAGCAGATGAGATACC	430
347 GTACCAAAACCTGAGCAAGATGAGGCTGAGAAAGTTCACTGAGCAGATGAGATACC	406
431 CCACTGGAGAGGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGAAATTCATAGATAGC	490
407 CCACTGGAGAGGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGAAATTCATAGATAGC	466
491 CATCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCATTCACAGAAATGCT	550
467 TATTTCTGAGCTTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCATTCACAGAAATGCT	526
551 GGTGGTAACTCTTTCATGATGATGCAAGAGAGGCTCTGAGAGCAAGAGAGAGATGT	610
527 GGTGGTAACTCTTTCATGATGATGCAAGAGAGGCTCTGAGAGCAAGAGAGAGATGT	586
611 TGTATTAAACAACCTTACATCACTCAAGAGAGGCTCAATTAACGAGAGAGGAGC	670

```

Db      587 TGTATTAAACGACGCTTACACTGCTATAGAGAAGACTGTGCACTACCGGGGTGACAGGAGC 646
Qy      671 GGTGAAGCTGCGCAAGGTTGGGGGCTTTGGCATCTCTCATTTGGATCCGTCCTCTCTC 730
Db      647 TGTGAAGCTGCGCAAGGTTGGGGGCTTTGGCATCTCTCATTTGGATCCGTCCTCTCTC 706
Qy      731 CATCTACAGCTCTGACACAGGATATTCAGAAATACAGAGATGGCTGCCAAATTCACAC 790
Db      707 CATCTACAGCTCTGACACAGGATATTCAGAAATACAGAGATGGCTGCCAAATTCACAC 766
Qy      791 AGCTGTATTAAGCTGTGAAGATGCAAAATATGATCAAGATGGCTCTCATGGATCAA 850
Db      767 AGCTGTATTAAGCTGTGAAGATGCAAAATATGATCAAGATGGCTCTCATGGATCAA 826
Qy      851 AATTGTATTAAGCTGTGAAGATGCAAAATATGATCAAGATGGCTCTCATGGATCAA 910
Db      827 AATTGTATTAAGCTGTGAAGATGCAAAATATGATCAAGATGGCTCTCATGGATCAA 886
Qy      911 TGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGCTAGTGTGACATGACATCT 970
Db      887 TGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGCTAGTGTGACATGACATCT 946
Qy      971 GGAAGCTGGAGATGTTGGGCGAGGCTGCATGATGATGCGGTGAGCTTTATATCATG 1030
Db      947 GGAAGCTGGAGATGTTGGGCGAGGCTGCATGATGATGCGGTGAGCTTTATATCATG 1006
Qy      1031 GGAAGCTCTCTCATTTTAAAGATCTTGGGCTGCTGCCAAGAGACTCTGCGGCTGCT 1090
Db      1007 GGAAGCTCTCTCATTTTAAAGATCTTGGGCTGCTGCCAAGAGACTCTGCGGCTGCT 1066
Qy      1091 GCTCTGACCTGACAGAGAACAGGTTGAGTGTGCTCTTCCAGATTTATGATACACAA 1150
Db      1067 GCTCTGACCTGACAGAGAACAGGTTGAGTGTGCTCTTCCAGATTTATGATACATTA 1126
Qy      1151 GGTAAATATTTTCCAACTACAGTCTGTGTGATGATGATGATGATGATGATGATGATGAT 1210
Db      1127 GGTAAATATTTTCCAACTACAGTCTGTGTGATGATGATGATGATGATGATGATGATGAT 1186
Qy      1211 TGGGCTGCAATTCATGCTGCAATGCAAAAGGCGGCGCATCTGAGAGAGGTTATGAGCT 1270
Db      1187 TGGGCTGCAATTCATGCTGCAATGCAAAAGGCGGCGCATCTGAGAGAGGTTATGAGCT 1246
Qy      1271 GCTGACAGCTCTCATATATCACTCAGGCTCTGAGCATGAGAGAGGACATCACTT 1330
Db      1247 TCTGCAACCTCTCATATGTCACCAAGCTCTTATGATGATGATGATGATGATGATGATGAT 1306
Qy      1331 TTGATCCAGCTGAGAGTGTGAGCGGCTGAGCGGCTGATCTGATGATGATGATGATGATGAT 1390
Db      1307 CTGATCCAGCTGAGAGTGTGAGCGGCTGAGCGGCTGATCTGATGATGATGATGATGATGAT 1366
Qy      1391 CTTTCATCACTCCACCGAGACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1450
Db      1367 TTTTCATCACTCCACCGAGACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426
Qy      1451 TGTGCTGTGTGGGCTGTGTTCTTATGTTGTCAGACATGGAAGAAATGTCGCTAG 1510
Db      1427 TGTGCTGTGTGGGCTGTGTTCTTATGTTGTCAGACATGGAAGAAATGTCGCTAG 1486
Qy      1511 GTCTAGAAACAGTAAGAAAGAAAGTTCATGCTTGTGCGCAGAAATCTGCGTCTGC 1570
Db      1487 GTCTAGAAACAGTAAGAAAGAAAGTTCATGCTTGTGCGCAGAAATCTGCGTCTGC 1546
Qy      1571 AACTTTGAAAACCTCTTTCATACATTAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 1630
Db      1547 ATCTTTACACATCTCATCTGACAAAGAGTGTATACATTTTCAACAGGCGATGTT 1606
Qy      1631 CATTTTCACTGCTTCTGTTATTTATTTCTTGTGATCTTCAAAATTTCTGATTTAGAA 1690
Db      1607 TTTCTTTATACCTTCTGTTATTTATTTCTTGTGATCTTCAAAATTTCTGATTTAGAA 1663
Qy      1691 AAGAATCATTTCTCCCTCCCTCC 1716
Db      1664 TAAATCATGATCCCTACTGACACACC 1689

```

```

RESULT 2
EC010977
LOCUS
DEFINITION
Mus musculus, similar to plasma glutamate carboxypeptidase, clone
IMAGE:4018296, mRNA.
ACCESSION
EC010977
VERSION
EC010977.1 GI:15012137
KEYWORDS
HTC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
Strauberg, R.
Direct Submission
Submitted (23-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcaps-x@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
Contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: http://image.liml.gov
Series: IRAC Plate: 18 Row: 9 Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 9055233
This clone has the following problem: retained intron.
FEATURES
Location/Qualifiers
source
1..1868
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4018296"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clone) outgrowth infected with the virus MMTV."
/clone_id="NCI CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
BASE COUNT
505 a 418 c 460 g 485 t
ORIGIN
Query Match 62.9%; Score 1164; DB 11; Length 1868;
Best Local Similarity 83.5%; Pred. No. 4.3e-211;
Matches 1333; Conservative 0; Mismatches 260; Indels 3; Gaps 1;
Qy 121 TTTTGGTGTGTTACCTTTTATCCCTGCTGTGGAAGATATGCAAGATGGCA 180
Db 217 TTTCTGTTATGTTACCTTTTATGCTGTGTAAGCTGTATGCAAGATGGT 276
Qy 181 TCTTAAAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTTCTA 240
Db 277 TTTCTAGCAACATTTTCGAGAAATAAAGAAATAGCCACTATGAGATGTTCTA 336
Qy 241 AAGCAATCAACCTGCTTTATGTTAAAGCCGCAACATCTTATGAGGATTTGG 300
Db 337 AAGCAATTAACCTTGTGTTATGTTAAATACGAAACGGTCTTATGAGGATTTGG 396
Qy 301 CACTTGTGTTATGCTGTGACCCAGATGAGTGTCTCAAGAACTTGAAGAAAGCA 360

```

Db 397 GACTTCTAGATTGATCTGTGGAGCCAGCTGAGTGGCTTAAAGAACTAGAGAAAGCTA 456
 Qy 361 TCCAAATTAATGTAACAAAACCTGACGAAAGATGGCTGGAGAAATTCACCTGGAGCCAG 420
 Db 457 TCCAAATATGTAACAAAACCTGACGAAAGATGGCTGGAGAAATTCACCTGGAGCCAG 516
 Qy 421 TGAGATATACCCCACTGGAGAGAGAGAAAGATTCAGCTGTGTGATGCTGGAGCCAAATTC 480
 Db 517 TCAGAAATACCCCACTGGAGAGAGAGAAAGATTCAGCTGTGTGATGCTGGAGCCAAATTC 576
 Qy 481 ATAAATATGCAATCTGGAGCTGGAGAGAGATTCAGCTGTGTGATGCTGGAGCCAAATTC 540
 Db 577 ACAATATGCTATCTTGTAGCTGGAGAGAGATTCAGCTGTGTGATGCTGGAGCCAAATTC 636
 Qy 541 CAGAGTCTGGTGTGAGCTCTTGTGATGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 Db 637 CAGAGTCTGGTGTGAGCTCTTGTGATGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 696
 Qy 601 GGAAGATTTGTTTATTAACCAACCTTACATCACTCACTCAAGAGAGAGAGAGAGAGAG 660
 Db 697 GGAAGATCTGTTTATTAACCAACCTTACATCACTCACTCAAGAGAGAGAGAGAGAGAG 756
 Qy 661 GCGAG 720
 Db 757 TGCAAG 816
 Qy 721 CCTCTTCTCCATCTACAGCTCTCTCAACAGATTCAGAGAAATTCAGAGAGAGAGAGAGAG 780
 Db 817 CT 876
 Qy 781 AAATTCCAACAGAGCTGTATTAACGAGTGAAGATTCAGAGAAATTCAGAGAAATTCAGAG 840
 Db 877 AGATTCCAACAGAGCTGTATTAACGAGTGAAGATTCAGAGAAATTCAGAGAAATTCAGAG 936
 Qy 841 ATGGATCAAAATTTGTCACTTACGCTTAAAGATGGAGAGAGAGAGAGAGAGAGAGAGAG 900
 Db 937 GTGGAGAAAGAAATTTGTCACTTACGAGTGAAGATGGAGAGAGAGAGAGAGAGAGAGAG 996
 Qy 901 CCTTCAACAGCTGTACAGAGATCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 Db 997 CCTTCAACAGCTGTACAGAGATCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1056
 Qy 961 GTGACATCTGGAGAGCTGGAGATGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 Db 1057 GTGACATCTGGAGAGCTGGAGATGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1116
 Qy 1021 TTATATCATGGAGAGAGCTCTCACTTATTAAGATCTTGGAGCTGTCTCAAGAGAGAGAG 1080
 Db 1117 TCATATCATGGAGAGAGCTCTCACTTATTAAGATCTTGGAGCTGTCTCAAGAGAGAGAG 1176
 Qy 1081 TGCGGCTGTGTCTGTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 Db 1177 TGCGGCTGTGTCTGTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1236
 Qy 1141 AGTTACACAGAGTAATATTTTCACTACAGCTGTGTGATGAGAGAGAGAGAGAGAGAGAG 1200
 Db 1237 AGTTACACAGAGTAATATTTTCACTACAGCTGTGTGATGAGAGAGAGAGAGAGAGAGAG 1296
 Qy 1201 TCTTACCACTGGAGCTGAGATTCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 Db 1297 TCTTACCACTGGAGCTGAGATTCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1356
 Qy 1261 TTATAGAGCTGTGAG 1320
 Db 1357 TCATATCATTTGTGAG 1416
 Qy 1321 ACATCAACTTTTGTGAG 1380
 Db 1417 ATATATCACTTGTGAG 1476
 Qy 1381 AGTATTTCTTCTTCATCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 Db 1477 AGTATTTCTTCTTCATCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1536

Qy 1441 TGAATGTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1500
 Db 1337 TGAATGTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1556
 Qy 1501 TGCTGCTGAGTCTGTGAG 1560
 Db 1597 TGTTGCCAGAGTCTGTGAG 1656
 Qy 1561 CTGGGCTGTGAG 1620
 Db 1657 CCAACTGTGAG 1716
 Qy 1621 AAGCAACTGTATTTTATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1680
 Db 1717 GGGATAGTCTTCTTATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1773
 Qy 1681 GATTTAG 1716
 Db 1774 GTTCTAG 1809
 RESULT 3
 AK075686
 LOCUS
 DEFINITION
 Mus musculus 1805 bp mRNA linear HTC 07-DEC-2002
 enriched library, clone:1190003p12 product:plasma glutamate
 carboxypeptidase, full insert sequence.
 ACCESSION
 AK075686
 VERSION
 AK075686.1 GI:26344480
 KEYWORDS
 HTC, CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 20499374
 11042159
 REFERENCE
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 11076861
 REFERENCE
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
 Komno, H., Akiyama, T., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishogi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Iwama, H., Ohara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multiplexed sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 20530913
 11076861
 REFERENCE
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pezole, G.,
 Quackenbush, J., Schiml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Baren, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bulic, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hotmann, M., Hume, D.A., Kamuya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Saeki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Tयो-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Borja, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseki, S. and Hayashizaki, Y.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE 5

AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 1805)

AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arahawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komoto, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaoka, I., Yasunaka, Y., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-ARR-2002) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES Please visit our web site for further details.

SOURCE URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
URL: http://location/Qualifiers

1..1805

location="Mus musculus"

organism="Mus musculus"

molecule="mRNA"

strain="C57BL/6J"

db_xref="PANTOM DB:1190003P12"

db_xref="MGI:1909852"

db_xref="taxon:10090"

clone="1190003P12"

tissue_type="whole body"

clone_lib="RIKEN full-length enriched mouse cDNA library"

dev_stage="18-day embryo"

147..1559

note="unlabeled protein product; plasma glutamate carboxypeptidase (MCP|MGI:1889205, GBLNM_018755, evidence: BLASTN, 99%, match=943)

putative"

codon_start=1

protein_id="BAC35891.1"

db_xref="MGI:26344481"

db_xref="MGI:1923585"

translation="MRSLLFLFLFYHLALSGSKAVFKNGVSQRFREIKETIANEDVAKAIINLAVGKYONRSYERLGLVDVTPPLSGSKLEKAIQMYNLODGLNVHLEOVYPMWEGESAVMLBPRHMKALIGGSSITGPGGITAEVNVAVSFDELQRLASEARGLIIVNPTGYEKTVOYRGVAGVAAKAVAVASLIGSVASFSIYSPHTGIO

KYDGVPKIPIPTACTIVEDAEEMSRMSRKNKIVIHLEMGAKYPTDTSFNTVAETIGS
MYEEFVLVSQDSDVVGALDDGGAFISWEALSLVDGLRKRTRILVLTAE
BOGGEIAYOYEEHKNISKYSILMEADSGTFPLPTGLOFGSKAKAIKEVNNLOP
LNTKXFSNEEGDINFMIOAGVPGASLDDLYKFFPHSHSDMTVMDDPKOMNVAA
AVAAVAVYVYVADMDENLPRS"

polya_signal
1788..1793
/note="putative"

polya_site
1805
/note="putative"

BASE COUNT 486 a 397 c 445 g 476 t 1 others

ORIGIN

Query Match 62.8%; Score 1162.2; DB 11; Length 1805;
Best Local Similarity 83.7%; Pred. No. 9.5e-211;
Matches 1327; Conservative 0; Mismatches 256; Indels 3; Gaps 1;

QY 131 TGTTCACCTTTTATCCCTGTGCTCTGCGGAAAGCTAATGCAAGAATGCACTCTTAAGAG 190
|||
Db 173 TGTTCACCTTTTATCCCTGTGCTCTGCGGAAAGCTAATGCAAGAATGCTTTCTCAGCG 232

QY 191 GACTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTTGCTAAAGCAATCAT 250
|||
Db 233 AACATTGAGAAATAAAGAAATAGCCAGCTAATGATGTTGCTAAAGCAATCAT 292

QY 251 CAACCTAGCTGTTTATGTTAAAGCCAGAAAGATCCTATGAGCAATGCACTTCTGT 310
|||
Db 293 CAACCTAGCTGTTTATGTTAAATACCAAGACCGGTCTATGAGCGTTTGGACCTTCACT 352

QY 311 TGATACCTGTTGAAGCCAGACTGATGCTGCTCAAGAACTTGAAGAAAGCAATTCAT 370
|||
Db 353 TGAATACCTGTTGAAGCCAGACTGATGCTGCTCAAGAACTTGAAGAAAGCAATTCAT 412

QY 371 GTACCAAAACCTGAGAAATAGGCGTGGAGAAAGTTCACTGAGCCATGAGAAATACC 430
|||
Db 413 GTACCAAAACCTGAGAAATAGGCGTGGAGAAAGTTCACTGAGCCATGAGAAATACC 472

QY 431 CCACCTGAGAGAGGAGAGAAATCAGCTGATGCTGAGCCCAAGAAATCATTAAGATAGC 490
|||
Db 473 CCACCTGAGAGAGGAGAGAAATCAGCTGATGCTGAGCCCTGAAATTCAGAAATAGGC 532

QY 491 CATCTGGGCTTTGGCAGCAGCAATGGGACTCTCCAGAAAGCAATTCAGCAGAAATGCT 550
|||
Db 533 TATCTGAGCTTTGGCAGCAGCAATGGGACTCTCCAGAAAGCAATTCAGCAGAAATGCT 592

QY 551 GGTGCTGACCTCTTTCATGATTAATCGCAGAGAAAGGCTCCAGAAAGCAAGAGAAATGT 610
|||
Db 593 GGTGCTGACCTCTTTCATGATTAATCAAGAGCAATTCAGAAAGCAAGAGAAATCAT 652

QY 611 TGTTTATACCAACCTTACATCACTCAAGAGAGGTCGCAATACGCAAGAGAGGAGG 670
|||
Db 653 TGTTTATACCAACCTTACATCACTGAGCTTGAAGAGCTGTCAGTACCGGCTGAGGAGG 712

QY 671 GGTGAGAGCTGCAAGAGTGGGAGCTTTGGCATCTTCATTCGATTCCTGAGCTCTCTTC 730
|||
Db 713 TGTGAGAGCTGCAAGAGTGGGAGCTTTGGCATCTTCATTCGATTCCTGAGCTCTCTTC 772

QY 731 CATTTACAGTCTCAGCAGAGTATTCAGAAATACAGAGAGGCTGCCAAATTCACAC 790
|||
Db 773 CATTTACAGTCTCAGCAGAGTATTCAGAAATACAGAGAGGCTGCCAAATTCACAC 832

QY 791 AGCTGATTTACGGTGGAAATGATGAGAAATGATGCAAGATGGCTCTCATGGGATCAA 850
|||
Db 833 AGCTGATTTACGGTGGAAATGATGAGAAATGATGCAAGATGGCTCTCATGGGATCAA 892

QY 851 AATTGCTATTCAGTAAAGATGGGAGCAAGACCTTACCCAGATATGATTCCTTCAACAC 910
|||
Db 893 AATTGCTATTCAGTAAAGATGGGAGCAAGACCTTACCCAGATATGATTCCTTCAACAC 952

QY 911 TGTGAGAGATTCAGTGGAGCAATATTCAGAAAGAGTGTACTGCTAGTGCACATCT 970
|||
Db 953 TGTGAGAGATTCAGTGGAGCAATATTCAGAAAGAGTGTCTGCTAGTGCACATCT 1012

QY 971 GGAAGCTGGAGATGTTGGAGAGGCTCCATGATGATGAGGCGGTGAGCCTTTATATCATG 1030

```

Db      1013  GGACAGCTGGAGTGTGGGCGAGGTCACATGATGATGCTGGAGCCTTCATATCATG 1072
Qy      1031  GGAAGCAGCTCTCACTATTAAGATCTTGGGCTGGCTCCAAAGAGAGACTCGGCTGCT 1090
Db      1073  GGAAGCAGCTCTCACTATTAAGATCTTGGGCTGGCTCCAAAGAGAGACTCGGCTGCT 1132
Qy      1091  GCTCTGAGCTGAGAAAGCAAGAGTGGATGGTCTCCAGTATTAATCACTTACACA 1150
Db      1133  GCTCTGAGCTGAGAAAGCAAGAGTGGATGGTCTCCAGTATTAATCACTTACACA 1192
Qy      1151  GGTAAATATTTCCAACTACAGTCTGGTGAAGTCTGACGAGAACTTCTTACCCAC 1210
Db      1193  GGTAAATATTTCCAACTACAGTCTGGTGAAGTCTGACGAGAACTTCTTACCCAC 1252
Qy      1211  TGGGCTGCAATTCACCTGAGAGAAAGGCGGCGCATCATGAGAGAGTATAGAGCT 1270
Db      1253  TGGAGCTGAGTTCACCTGAGAGAAAGGCGGCGCATCATGAGAGAGTATAGAGCT 1312
Qy      1271  GCTGAGAGCTCTCAATATCACTCAGGCTCTGAGCCATGAGAGAGGACATCACTT 1330
Db      1313  TCTGCAAGCTCTCAATATCACTCAGGCTCTGAGCCATGAGAGAGTATAGAGCTT 1372
Qy      1331  TTGATTCAGAGCTGAGAGTCTGAGAGGCTCTGAGCCATGAGAGAGTATAGAGCTT 1390
Db      1373  CTGATTCAGAGCTGAGAGTCTGAGAGGCTCTGAGCCATGAGAGAGTATAGAGCTT 1432
Qy      1391  CTTCATCATCTCCACGAGAGACATGATCTGATGATCCAAAGAGAGAGTATAGAGCT 1450
Db      1433  TTTCATCATCTCCACGAGAGACATGATCTGATGATCCAAAGAGAGAGTATAGAGCT 1492
Qy      1451  TGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGAGAAATGCTGCTAG 1510
Db      1493  TGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGAGAAATGCTGCTAG 1552
Qy      1511  GTCTGAGAAAGAGTAAAGAAAGAGTTCATGCTTCTGCGCAGAGATCTCGGCTGC 1570
Db      1553  GTCTGAGAGAGAAAGAGAGAGAGAGCTTCTGCTGTAAGTGGAAATCCCACTGTA 1612
Qy      1571  AACTTGGAGAAAGCTCTCTTCAATACAAATTCATTCATCTTCAAGAGCAACT 1630
Db      1613  AACTTGGAGAAAGCTCTCTTCAATACAAATTCATTCATCTTCAAGAGCAACT 1672
Qy      1631  CTAATTCAGCTCTTCTGTTATATCTTCTGATCTTCAATCTTCTGATCTTCAAG 1690
Db      1673  TTTCTTATACCTTCTGTTATATCTTCTGATCTTCAATCTTCTGATCTTCAAG 1729
Qy      1691  AAAAGATCATTTCTCCCTCCCTCC 1716
Db      1730  TAAATCATGATCCCTACTGACACACC 1755

RESULT 4
BC012019      1024 bp      mRNA      linear      HTC 04-MAR-2003
LOCUS      Homo sapiens, Similar to plasma glutamate carboxypeptidase, clone
DEFINITION      IMAGE:4455631, mRNA.
ACCESSION      BC012019
VERSION      BC012019.1
KEYWORDS      GI:15080558
SOURCE      HTC.
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 1024)
AUTHORS      Strausberg, R.
TITLE      Direct Submission
JOURNAL      Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk

```

```

FEATURES
    source
        1..1024
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4455631"
            /tissue_type="kidney, hypernephroma"
            /clone_lib="NIH MGC_89"
            /lab_host="DH10B"
            /note="Vector: pCMV-SPORT6"

BASE COUNT      317 a      210 c      256 g      241 t
ORIGIN
Query Match      53.7%; Score 993.8; DB 11; Length 1024;
Best Local Similarity 99.8%; Pred. No. 1,1e-178;
Matches 995; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

64 GATTATCTTAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCGATTTT 123
1 GATTATCTTAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCGATTTT 60
Qy      124  TCGGTGTTACCTTTTATCCCTGCTCTGCGAAAGCTATATGCAAGATGCAATCT 183
Db      61  TCGGTGTTACCTTTTATCCCTGCTCTGCGAAAGCTATATGCAAGATGCAATCT 120
Qy      184  CTAAGAGAGCTTTGAAGAAATTAAGAAATAGCAGTGTGAGATGCTTAAG 243
Db      121  CTAAGAGAGCTTTGAAGAAATTAAGAAATAGCAGTGTGAGATGCTTAAG 180
Qy      244  CAATCATCAACTAGCTGTTATGTTAAAGCCAGAAAGATCTTATGAGCGATTGGCAC 303
Db      181  CAATCATCAACTAGCTGTTATGTTAAAGCCAGAAAGATCTTATGAGCGATTGGCAC 240
Qy      304  TTCTGTTGATCTGTTGAGCCAGACTGAGTGTCTCAAGAACTTGAAGAAAGCCATCC 363
Db      241  TTCTGTTGATCTGTTGAGCCAGACTGAGTGTCTCAAGAACTTGAAGAAAGCCATCC 300
Qy      364  AAATTAATGTAACCAAAACCTGAGCAAGATAGGCTGGAAGAAAGTTCACTGAGCACTGA 423
Db      301  AAATTAATGTAACCAAAACCTGAGCAAGATAGGCTGGAAGAAAGTTCACTGAGCACTGA 360
Qy      424  GAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGATGCTGAGACCAAGATTCATA 483
Db      361  GAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGATGCTGAGACCAAGATTCATA 420
Qy      484  AGATAGCCATCTGGGTTCTTGGACAGCAATGGGACTCTCCAGAAAGCATTAACAGCAG 543
Db      421  AGATAGCCATCTGGGTTCTTGGACAGCAATGGGACTCTCCAGAAAGCATTAACAGCAG 480
Qy      544  AAGTTCTGAGTGAACCTTTTGGATGATCACTGAGAGAGGCTCTGAGAAAGAGGA 603
Db      481  AAGTTCTGAGTGAACCTTTTGGATGATCACTGAGAGAGGCTCTGAGAAAGAGGA 540
Qy      604  AGATTGTTGTTTAAACCAACTTACATCAACTACTAGAGAGGCTGCAATACCGAAGC 663

```

Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amebcm.tmc.edu
Gunsberg, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisse, H.,
Kovis, C.R., Sneed, A.O., Martin, R.G., Muzny, D.W., Nantavali,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAC Plate: 28 Row: b Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5174626
This clone has the following problem: retained intron.

FEATURES

source

BASE COUNT

317 a 210 c 256 g 241 t

ORIGIN

Query Match 53.7%; Score 993.8; DB 11; Length 1024;

Best Local Similarity 99.8%; Pred. No. 1,1e-178; Matches 995; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

64 GATTATCTTAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCGATTTT 123
1 GATTATCTTAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCGATTTT 60
Qy      124  TCGGTGTTACCTTTTATCCCTGCTCTGCGAAAGCTATATGCAAGATGCAATCT 183
Db      61  TCGGTGTTACCTTTTATCCCTGCTCTGCGAAAGCTATATGCAAGATGCAATCT 120
Qy      184  CTAAGAGAGCTTTGAAGAAATTAAGAAATAGCAGTGTGAGATGCTTAAG 243
Db      121  CTAAGAGAGCTTTGAAGAAATTAAGAAATAGCAGTGTGAGATGCTTAAG 180
Qy      244  CAATCATCAACTAGCTGTTATGTTAAAGCCAGAAAGATCTTATGAGCGATTGGCAC 303
Db      181  CAATCATCAACTAGCTGTTATGTTAAAGCCAGAAAGATCTTATGAGCGATTGGCAC 240
Qy      304  TTCTGTTGATCTGTTGAGCCAGACTGAGTGTCTCAAGAACTTGAAGAAAGCCATCC 363
Db      241  TTCTGTTGATCTGTTGAGCCAGACTGAGTGTCTCAAGAACTTGAAGAAAGCCATCC 300
Qy      364  AAATTAATGTAACCAAAACCTGAGCAAGATAGGCTGGAAGAAAGTTCACTGAGCACTGA 423
Db      301  AAATTAATGTAACCAAAACCTGAGCAAGATAGGCTGGAAGAAAGTTCACTGAGCACTGA 360
Qy      424  GAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGATGCTGAGACCAAGATTCATA 483
Db      361  GAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGATGCTGAGACCAAGATTCATA 420
Qy      484  AGATAGCCATCTGGGTTCTTGGACAGCAATGGGACTCTCCAGAAAGCATTAACAGCAG 543
Db      421  AGATAGCCATCTGGGTTCTTGGACAGCAATGGGACTCTCCAGAAAGCATTAACAGCAG 480
Qy      544  AAGTTCTGAGTGAACCTTTTGGATGATCACTGAGAGAGGCTCTGAGAAAGAGGA 603
Db      481  AAGTTCTGAGTGAACCTTTTGGATGATCACTGAGAGAGGCTCTGAGAAAGAGGA 540
Qy      604  AGATTGTTGTTTAAACCAACTTACATCAACTACTAGAGAGGCTGCAATACCGAAGC 663

```

```

Db      541 AGATTGTTTATTAACCACTTACCACTACTCAAGGACGGTGCATAATCCGAACGC 600
Qy      664 AGGGGGGGGTGAGAGCTGCCAAGGTGGGGGCTTTGGCATCTCTATTCGATCCGTGGCCT 723
Db      601 AGGGGGGGGTGAGAGCTGCCAAGGTGGGGGCTTTGGCATCTCTATTCGATCCGTGGCCT 660
Qy      724 CCTTCTCATCTACAGTCTCCACACAGGTATTCAGAAATACAGAGATGGCGCCCAAAA 783
Db      661 CCTTCTCATCTACAGTCTCCACACAGGTATTCAGAAATACAGAGATGGCGCCCAAGA 720
Qy      784 TTCCAACAGCTGTATTACGGTGGAGATGCAAAATGATGTCAAAATGGCTTTCATG 843
Db      721 TTCCAACAGCTGTATTACGGTGGAGATGCAAAATGATGTCAAAATGGCTTTCATG 780
Qy      844 GGATCAAAATGTCTATTCAGTAAAGATGGGGGCAAAAGCTTACCAGATCTGATTCCT 903
Db      781 GGATCAAAATGTCTATTCAGTAAAGATGGGGGCAAAAGCTTACCAGATCTGATTCCT 840
Qy      904 TCAACACTGTAGAGAGATCACTGGAGCAATATCCAGACAGGTTGTACTGGTCAAGT 963
Db      841 TCAACACTGTAGAGAGATCACTGGAGCAATATCCAGACAGGTTGTACTGGTCAAGT 900
Qy      964 GACATCTGACAGCTGGAGTGTGGGCAAGGTGCCATGATGATGGCGGTGAGCCTTTA 1023
Db      901 GACATCTGACAGCTGGAGTGTGGGCAAGGTGCCATGATGATGGCGGTGAGCCTTTA 960
Qy      1024 TATCATGGGAAGCACTCTCACTTATTAAGATCTTTGG 1060
Db      961 TATCATGGGAAGCACTCTCACTTATTAAGATCTTTGG 997

```

RESULT 5

```

BX355939/c 996 bp mRNA linear EST 05-MAY-2003
LOCUS BX355939 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CSODI006Y115 3-PRIME, mRNA sequence.
ACCESSION BX355939
VERSION BX355939.1 GI:30382027
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 996)
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODI006AB08&picluster=4663.f. Contact :
Peng Liang Email : liliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI006AB08NP1.
Location/Qualifiers
1..996
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI006Y115"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and BclI V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 270 a 231 c 223 g 269 t 3 others

```

ORIGIN

Query Match 53.64; Score 992.8; DB 13; Length 996;
 Best Local Similarity 99.64; Pred. No. 1.7e-178;
 Matches 992; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

Qy      788 AACAGCCTGTATTACCGTGGAGATGCAAAATGATGTCAAGATGGCTTCTCATGGAT 847
Db      996 AACAGCCTGTATTACCGTGGAGATGCAAAATGATGTCAAGATGGCTTCTCATGGAT 937
Qy      848 CAAATTTGATTCAGCTTAAAGATGGGGGCAAAAGACTTACCAGATTCGATTCCTTCA 907
Db      936 CAAATTTGATTCAGCTTAAAGATGGGGGCAAAAGACTTACCAGATTCGATTCCTTCA 877
Qy      908 CACTGTAGCAGATGTCCTGGAGCAAAATTCAGAAACAGGTGTACTGGTCAAGTGA 967
Db      876 CACTGTAGCAGATGTCCTGGAGCAAAATTCAGAAACAGGTGTACTGGTCAAGTGA 817
Qy      968 TCTGACAGCTGGAGATTTGGGAGAGGTGCCATGATGATGGCGGTGAGCCTTTATATC 1027
Db      816 TCTGACAGCTGGAGATTTGGGAGAGGTGCCATGATGATGGCGGTGAGCCTTTATATC 757
Qy      1028 ATGGGAAGCACTCTCACTTATTAAGATCTTTGGGCTGCTCAAGAGACTTGGCGCT 1087
Db      756 ATGGGAAGCACTCTCACTTATTAAGATCTTTGGGCTGCTCAAGAGACTTGGCGCT 697
Qy      1088 GGTGCTCTGACATGCAAGAAAGCAAGGTGAGTTGGTCCCTTCAGATTAATACGTTACA 1147
Db      696 GGTGCTCTGACATGCAAGAAAGCAAGGTGAGTTGGTCCCTTCAGATTAATACGTTACA 637
Qy      1148 CAAGGTAAATATTTCACACTACAGTCTGTGATGATGATGATGATGATGATGATGATG 1207
Db      636 CAAGGTAAATATTTCACACTACAGTCTGTGATGATGATGATGATGATGATGATGATG 577
Qy      1208 CACTGGCTGCAATTCACCTGAGTGAAGAAAGCCAGGCGCATCATGAGAGGTTATGAG 1267
Db      576 CACTGGCTGCAATTCACCTGAGTGAAGAAAGCCAGGCGCATCATGAGAGGTTATGAG 517
Qy      1268 CCGTGTGACAGCCCTCAATATCACTGAGTCTGAGGCAATGAGAAAGGACAGCATCAA 1327
Db      516 CCGTGTGACAGCCCTCAATATCACTGAGTCTGAGGCAATGAGAAAGGACAGCATCAA 457
Qy      1328 CTTTGGATCCAAGCTGAGTGGCTGGAGCAGTCTCTTATGATGATGATGATGATGATG 1387
Db      456 CTTTGGATCCAAGCTGAGTGGCTGGAGCAGTCTCTTATGATGATGATGATGATGATG 397
Qy      1388 CTTTCCATCACTCCACGAGAGACACCATGACTGTGATGATGATGATGATGATGATG 1447
Db      396 CTTTCCATCACTCCACGAGAGACACCATGACTGTGATGATGATGATGATGATGATG 337
Qy      1448 TGCCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTCAGACATGGAAGAAATGCTGCC 1507
Db      336 TGCCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTCAGACATGGAAGAAATGCTGCC 277
Qy      1508 TAGTCTCTAGAAACAGTAGAAGAAAGTTTTCATGCTTGGCCAGAGATCCTGGGTC 1567
Db      276 TAGTCTCTAGAAACAGTAGAAGAAAGTTTTCATGCTTGGCCAGAGATCCTGGGTC 217
Qy      1568 TGCACCTTTGAAAACCTCTCTTCACTAACATTAATTCATCAATTCCTTCAAGACA 1627
Db      216 TGCACCTTTGAAAACCTCTCTTCACTAACATTAATTCATCAATTCCTTCAAGACA 157
Qy      1628 ACTGTATTTCAATGCTTCTGTATTAATGTTTCTTATGATGATGATGATGATGATG 1687
Db      156 ACTGTATTTCAATGCTTCTGTATTAATGTTTCTTATGATGATGATGATGATGATG 97
Qy      1688 GAAAAAGAAATCAATTCCTCCCTCCCTCCACACATGAAATCAATGATGATGATGATG 1747
Db      96 GAAAAAGAAATCAATTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 37
Qy      1748 CAGTGGGGGCAATTTCTTATATCACTCTTAAAAAC 1783
Db      36 CAGTGGGGGCAATTTCTTATATCACTCTTAAAAAC 1

```



```

RESULT 6
BX439467
LOCUS      BX439467 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE010YJ06
DEFINITION
5-PRIME, mRNA sequence.
ACCESSION  BX439467
VERSION     BX439467.1 GI:30777746
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE  1 (bases 1 to 1050)
AUTHORS    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished
COMMENT     Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE010DE030P1c1cluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE010DE030P1.
location/Qualifiers
1..1050
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE010YJ06"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT  307 a 233 c 271 g 239 t
ORIGIN
Query Match      53.3%; Score 985.8; DB 13; Length 1050;
Best Local Similarity 99.8%; Pred. No. 3,6e-177;
Matches 987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
19  GTCCGAGGAGCGCGCTCGCGCGCCGACAGCGCGCTTACGATTATCTTAACAG 78
    |||||||
62  GTCCGAGGAGCGCGCTCGCGCGCCGACAGCGCGCTTACGATTATCTTAACAG 121
    |||||||
79  AAACCACTGGAAGAAAAATGAATTCCTATCTTCGATTTTCGCTGTTACAC 138
    |||||||
122  AAACCACTGGAAGAAAAATGAATTCCTATCTTCGATTTTCGCTGTTACAC 181
    |||||||
139  TTTTATCCCTGCTCTGCGAAGCTATATGCAAGATGCAATCTTAAGAGACTTTTG 198
    |||||||
182  TTTTATCCCTGCTCTGCGAAGCTATATGCAAGATGCAATCTTAAGAGACTTTTG 241
    |||||||
199  AAGAAATTAAGAAAGAAATGAGCACTGTGAGATGTTGCTTAAGCAATCATCACTAG 258
    |||||||
242  AAGAAATTAAGAAAGAAATGAGCACTGTGAGATGTTGCTTAAGCAATCATCACTAG 301
    |||||||
259  CTGTTTATGATGAAGCCGAGAACATCTATGACGATTTGCACTTCCTGTTGATACG 318
    |||||||
302  CTGTTTATGATGAAGCCGAGAACATCTATGACGATTTGCACTTCCTGTTGATACG 361
    |||||||
319  TTGACCCAGACTGAGTGGCTTCAAGAACCTTAAGAAAAAGCCATCCAAATTAATGACCAA 378
    |||||||
362  TTGACCCAGACTGAGTGGCTTCAAGAACCTTAAGAAAAAGCCATCCAAATTAATGACCAA 421
    |||||||

```

```

379  ACTGCAGCAAGATGGGCTGGAGAAAGTTCACTGAGAGCCAGTGAATACCCCACTGGG 438
    |||||||
422  ACTGCAGCAAGATGGGCTGGAGAAAGTTCACTGAGAGCCAGTGAATACCCCACTGGG 481
    |||||||
439  AGAGGGGAGAAAGATTCAGCTGTGTGATGCTGAGAGCCAGAAATTCATTAAGTACCTCTGG 498
    |||||||
482  AGAGGGGAGAAAGATTCAGCTGTGTGATGCTGAGAGCCAGAAATTCATTAAGTACCTCTGG 541
    |||||||
499  GTCTTGGCAGAGATTTGGGACTCTCTCCAGAAAGCAATTAACAGCAAGTTCTGTGTGTA 558
    |||||||
542  GTCTTGGCAGAGATTTGGGACTCTCTCCAGAAAGCAATTAACAGCAAGTTCTGTGTGTA 601
    |||||||
559  CCTCTTTGATGAATCTGAGAGAGGCTCTCAGAGCAAGAGGAGAAATTTGTTTATA 618
    |||||||
602  CCTCTTTGATGAATCTGAGAGAGGCTCTCAGAGCAAGAGGAGAAATTTGTTTATA 661
    |||||||
619  ACCAACCTTACATCACTACTCAAGAGCGTGCAATACCGAACGAGGGGCGGTGGAAG 678
    |||||||
662  ACCAACCTTACATCACTACTCAAGAGCGGTGCAATACCGAACGAGGGGCGGTGGAAG 721
    |||||||
679  CTGCAAGGTGGGGCTTTGGCATCTCTCATTTGCATTCCTGCTCTTCTTCAATCTTACA 738
    |||||||
722  CTGCAAGGTGGGGCTTTGGCATCTCTCATTTGCATTCCTGCTCTTCTTCAATCTTACA 781
    |||||||
739  GTCCCTACACAGATTAATCAGGAATACAGAGATGGCGTCCCAAAATTCACACAGCTGTA 798
    |||||||
782  GTCCCTACACAGATTAATCAGGAATACAGAGATGGCGTCCCAAAATTCACACAGCTGTA 841
    |||||||
799  TTACGCTGAAGATGCAAGAAATGATGTCAGAAATGCTTCTCATGAGATCAAAATTTGCA 858
    |||||||
842  TTACGCTGAAGATGCAAGAAATGATGTCAGAAATGCTTCTCATGAGATCAAAATTTGCA 901
    |||||||
859  TTACGCTGAAGATGCAAGAAATGATGTCAGAAATGCTTCTCATGAGATCAAAATTTGCA 918
    |||||||
902  TTACGCTGAAGATGCAAGAAATGATGTCAGAAATGCTTCTCATGAGATCAAAATTTGCA 961
    |||||||
919  AGATCACTGGAGCAAAATATCAGAACAGGTTGACTGTCAGTGAACATCTGACAGACT 978
    |||||||
962  AGATCACTGGAGCAAAATATCAGAACAGGTTGACTGTCAGTGAACATCTGACAGACT 1021
    |||||||
979  GGAGTGTGGGCGAGGTGCCATGATGAT 1007
    |||||||
1022  GGAGTGTGGGCGAGGTGCCATGATGAT 1050
    |||||||

RESULT 7
BX355940
LOCUS      BX355940 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION
5-PRIME, mRNA sequence.
ACCESSION  BX355940
VERSION     BX355940.1 GI:30384019
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE  1 (bases 1 to 1038)
AUTHORS    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished
COMMENT     Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1006A080P1c1cluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

```


FEATURES
source

1. 1038

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D1006Y115"

/issue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 304 a 237 c 262 g 233 t 2 others

ORIGIN

Query Match 50.8%; Score 940.8; DB 13; Length 1038;
Best Local Similarity 99.7%; Pred. No. 1.3e-168;
Matches 953; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

2 GCTAGGCGGAGGCTTGTCTGAGAGCGGCTCGGCGCGGTACAGCCGCCCTAT 61
Db GCTAGGCGGAGGCTTGTCTGAGAGCGGCTCGGCGCGGTACAGCCGCCCTAT 142
QY CAGATTATCTTAAACAGAAACCAACTGGAAGAAATGAAATGCTTATCTTGCATT 121
Db CAGATTATCTTAAACAGAAACCAACTGGAAGAAATGAAATGCTTATCTTGCATT 202
QY 122 TTTGGTGTGTTCACCTTTTATCCCTGTGTCTGGAAGCTATATGCAAGATGCGAT 181
Db 203 TTTGGTGTGTTCACCTTTTATCCCTGTGTCTGGAAGCTATATGCAAGATGCGAT 262
QY 182 CTCTAAGAGGACTTTTGAAGAAATGAAAGAAATGCGAGCTGAGATGTTGCTAA 241
Db 263 CTCTAAGAGGACTTTTGAAGAAATGAAAGAAATGCGAGCTGAGATGTTGCTAA 322
QY 242 AGCAATCATCAACCTAGCTGTTTATGTTAAGCCAGAACGATCTATGAGCATGGC 301
Db 323 AGCAATCATCAACCTAGCTGTTTATGTTAAGCCAGAACGATCTATGAGCATGGC 382
QY 302 ACTTGTGTGTATCTGTGGAACCAAGCTAGTGGCTCCAGAACTTAGAAAAAGCCAT 361
Db 383 ACTTGTGTGTATCTGTGGAACCAAGCTAGTGGCTCCAGAACTTAGAAAAAGCCAT 442
QY 362 CCAATTTATGTAACCAAAACCTGACAGAAAGGGCTGAGAAAGTCCCTGGAAGCCAT 421
Db 443 CCAATTTATGTAACCAAAACCTGACAGAAAGGGCTGAGAAAGTCCCTGGAAGCCAT 502
QY 422 GAGAATACCCCACTGGAAGAGGAGGAAGATCACTGTGATGCTGAGCCAGAAATTC 481
Db 503 GAGAATACCCCACTGGAAGAGGAGGAAGATCACTGTGATGCTGAGCCAGAAATTC 562
QY 482 TAAGATAGCCATCTGTGGTCTTGGACAGCAGATTGGGATCTCTCCAGAAAGGCTTAC 541
Db 563 TAAGATAGCCATCTGTGGTCTTGGACAGCAGATTGGGATCTCTCCAGAAAGGCTTAC 622
QY 542 AGAAGTCTGTGTGTAACCTTTTGAAGAACTGCAAGAAAGGGCTTGAAGAAAGAG 601
Db 623 AGAAGTCTGTGTGTAACCTTTTGAAGAACTGCAAGAAAGGGCTTGAAGAAAGAG 682
QY 602 GAAGATTGTTGTTATATAACCACTTACATCAACTACAGAGAGCGTGCAATCCGAA 661
Db 683 GAAGATTGTTGTTATATAACCACTTACATCAACTACAGAGAGCGTGCAATCCGAA 742
QY 662 GCAGGGGCGGTGGAAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTTGATCCGTGG 721
Db 743 GCAGGGGCGGTGGAAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTTGATCCGTGG 802
QY 722 CTCTCTTCTTCATCTACAGTCTCTCAACAGGTATTCAGAAATACAGAGATGGCGCCAA 781
Db 803 CTCTCTTCTTCATCTACAGTCTCTCAACAGGTATTCAGAAATACAGAGATGGCGCCAA 862
QY 782 AATTCACAGCTGTATTAAGGTGGAAGATGCAAGAAATGATGTCAGAAATGGCTTCTCA 841

```

```

|||||
Db 863 GATTCACAGACCTGTATTAAGGTGGAAGATGCAAGAAATGATGCAAGAAATGGCTTCTCA 922
QY 842 TGGGATTAATAATGTCTATTGAGCTTAAAGAT-GGGGCGCAAGACCTTACCAGATTAAT 900
Db 923 TGGGATTAATAATGTCTATTGAGCTTAAAGATGGGGGCGCAAGACCTTACCAGATTAAT 982
QY 901 CCTTCAACACTGTAGCAGATATCACTGGAGCAATATTCAGAAACAGTTGTACTG 956
Db 983 CCTTCAACACTGTAGCAGATATCACTGGAGCAATATTCAGAAACAGTTGTACTG 1038

```

RESULT 8
LOCUS BX439466/c 1020 bp mRNA linear EST 15-MAY-2003
DEFINITION BX439466 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE010Y006
3-PRIME, mRNA sequence.
ACCESSION BX439466
VERSION BX439466.1 GI:30775754
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1020)
L.I.W.B., Gruber, C., Jesses, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE010DE03NP1&cluster=4663.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE010DE03NP1.

FEATURES
source

1. 1020

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DE010Y006"

/issue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 272 a 224 c 223 g 266 t 35 others

ORIGIN

Query Match 50.0%; Score 925.6; DB 13; Length 1020;
Best Local Similarity 95.9%; Pred. No. 9.9e-166;
Matches 925; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

```

850 AATTTGATTCACCTTAAGATGGGGCAAGACCTTCCAGATTAATGATCTTCTTACA 909
Db 966 AATTTGATTCACCTTAAGATGGGGCAAGACCTTCCAGATTAATGATCTTCTTACA 907
QY 910 CTGTAGCAGATATCACTGGAGCAAAATATTCAGAAACAGTTGTATCTGTGATGAGCATC 969
Db 906 CTGTAGCAGATATCACTGGAGCAAAATATTCAGAAACAGTTGTATCTGTGATGAGCATC 847
QY 970 TGGACAGCTGGATGTTGGGAGGGTCCATGATGATGATGCGCGGTGAGCCCTTATATCAT 1029
Db 846 TGGACAGCTGGATGTTGGGAGGGTCCATGATGATGATGCGCGGTGAGCCCTTATATCAT 787
QY 1030 GGGAGACATCTCACTTAATTAAGATCTTGGGGCTGCTCAAGAGAGACTCTGGGCTGG 1089

```


OY	1467	GTGTGTTCTTATGTTGTGCAAGACATGGAAGAAATGCTGCTAGGTCTTAGAAACAGTAA	1526
Db	330	GTGTGTTCTTATGTTGTGCAAGACATGGAAGAAATGCTGCTAGGTCTTAGAAACAGTAA	271
OY	1527	GAAGAAACGTTTTCATGCTTCTGGCCAGAAATCCTGGGTCTGCACATTTTGAAACCTCC	1586
Db	270	GAAGAAACGTTTTCATGCTTCTGGCCAGAAATCCTGGGTCTGCACATTTTGAAACCTCC	211
OY	1587	TCTTCACATTAACAATTTCATCCAAATTCATCTTCAAAAGACAACTGTATTTCAATGCTTCT	1646
Db	210	TCTTCACATTAACAATTTCATCCAAATTCATCTTCAAAAGACAACTGTATTTCAATGCTTCT	151
OY	1647	GTATATATCTTTCTTGATACCTTCCAAATTCCTGTATTTCTTGAAAAAGAGATCATTTCTCC	1706
Db	150	GTATATATCTTTCTTGATACCTTCCAAATTCCTGTATTTCTTGAAAAAGAGATCATTTCTCC	91
OY	1707	CCTCCCTCCACACATAGAAATCAACATAGGTAGAGATTAACAGTGGGGGACATTTCTTTTA	1766
Db	90	CCTCCCTCCACACATAGAAATCAACATAGGTAGAGATTAACAGTGGGGGACATTTCTTTTA	31
OY	1767	TATCACCTCTTAAAAACATTTGTTTCCACTT	1796
Db	30	TATCACCTCTTAAAAACATTTGTTTCCACTT	1

RESULT 10	1126 bp	EST 05-MAY-2003
EX360507		
LOCUS		
DEFINITION	1126 bp	EST 05-MAY-2003
DESCRIPTION	EX360507 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS001071Ya19 5-PRIME, mRNA sequence.	

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE 1 (bases 1 to 1126)

JOURNAL unpublished
COMMENT Contact: Genoscope

Genoscope - Centre National de Séquençage
RD 101 01006 Evry cedex, France

Email: segreff@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f For more information about this cluster, see

cg1-bln/cluster_cg17seq=CS0D1071AA10QPlcIcluster=4663.f. Contact:
Feng Liang Email: liang@lifetech.com URL: <http://fulllength.invitrogen.com/> INVITROGEN Corporation 1600
Paradise Avenue genoscope sequence ID: CS0D1071AA10QPl.

FEATURES

SOURCE

Location/Qualifiers
1. .1126

BASE COUNT	311 a	249 c	298 g	246 t	22 others
ORIGIN					

Query Match	47.1%;	Score 872.4;	DB 13;	Length 1126;
Best Local Similarity	97.0%;	Pred. No. 1.2e-155;		
Matches 931; Conservative	9;	Mismatches 15;	Indels 5;	Gaps 5.

OY	1	GGCTAGGCGCGGAGGCTTAGTCTGGGAGCGCCTCCCTCGCGGCTGACAGCGCGCCCTA	60
Db	112	GGCTAGGCGCGGAGGCTTAGTCTGGGAGCGCCTCCCTCGCGGCTGACAGCGCGCCCTA	171
OY	61	TCAGATTATCTTAAACAAGAAAACCACTGGAAAAAAAATGAAATTCCTTATCTTGGCAT	120
Db	172	TCAGATTATCTTAAACAAGAAAACCACTGGAAAAAAAATGAAATTCCTTATCTTGGCAT	231
OY	121	TTTTGGGTGGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATNTGCAAGATGGCA	180
Db	232	TTTTGGGTGGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATNTGCAAGATGGCA	291
OY	181	TCTCTAAGAGGACTTTTGAAGAAATTAAGAAGAAATAGCAGCTGTGGAGATTTGCTA	240
Db	292	TCTCTAAGAGGACTTTTGAAGAAATTAAGAAGAAATAGCAGCTGTGGAGATTTGCTA	351
OY	241	AAGCAATCATCAACTGACTGTTTATGTGTAAAGCCAGAAACAGATCCTATGAGCATTTG	300
Db	352	AAGCAATCATCAACTGACTGTTTATGTGTAAAGCCAGAAACAGATCCTATGAGCATTTG	411
OY	301	CACCTTCGTGTGATACGTGTGGACCCGACCTGAGTGGCTCCANAACCTTGAAAAACCA	360
Db	412	CACCTTCGTGTGATACGTGTGGACCCGACCTGAGTGGCTCCANAACCTTGAAAAACCA	471
OY	361	TCCAAATTAATGACCAAAACCTGAGAGAAATGGGCTGGAGAAAGTTCACTGGAGCGAG	420
Db	472	TCCAAATTAATGATACCAAAACCTGAGAGAAATGGGCTGGAGAAAGTTCACTGGAGCGAG	531
OY	421	TGAGAAATACCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATCTGGAGCCAGAAATTC	480
Db	532	TGAGAAATACCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATCTGGAGCCAGAAATTC	591
OY	481	ATAAGATAGCCATCTCTGGGCTTTGGGAGCAGCATTTGGGATCTCTCCAGAAAGGCAATTACG	540
Db	592	ATAAGATAGCCATCTCTGGGCTTTGGGAGCAGCATTTGGGATCTCTCCAGAAAGGCAATTACG	651
OY	541	CAGAAGTTCTGGTGGTGAACCTCTTCGATGAACTGCAAGAGAGGGGCTCGAAGCAAGAG	600
Db	652	CAGAAGTTCTGGTGGTGAACCTCTTCGATGAACTGCAAGAGAGGGGCTCGAAGCAAGAG	711
OY	601	GGAAGATTGTTGTTTATTAACCAACCTTAACATCACTA	660
Db	712	GGAAGATTGTTGTTTATTAACCAACCTTAACATCACTA	771
OY	661	CGCAGGGGGCGGTGGAAAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCATTTGCATTCGTTG	720
Db	772	CGCAGGGGGCGGTGGAAAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCATTTGCATTCGTTG	831
OY	721	CCTCCTTCGCATCTACAGTCTCAACAGGTAATTCAGGAATACAGGATGGCGTGCCCA	780
Db	832	CCTCCTTCGCATCTACAGTCTCAACAGGTAATTCAGGAATACAGGATGGCGTGCCCA	891
OY	781	AAATTTCCAAACGCTGTATTACGCT -GGAAGATGCAAGAAATGATGTCAAGATGGCTTCT	839
Db	892	AAATTTCCAAACGCTGTATTACGCTGTGGGAGAGATCAGAAATGATGTCAAGATGGCTTCT	951
OY	840	CA-TGGGATCAAAATTTGTCATTACGCTTAAGATGGGGGCAAAAGCTTAACCAATACTGA	898
Db	952	CATTGGGATCAAAATTTGTCATTACGCTTAAGATGGGGG -AARMACTAACCCAGATACTGA	1010
OY	899	TTCCCTTCAACCTGTAGAGAGATCACTGGGAGCAAAATATCCAAAGCTGTATACGCT	958
Db	1011	TTCCCTTCAACCTGTAGAGAGATCACTGGGAGCAAAATATCCAAAGCTGTATACGCT	1068

RESULT	11
LOCUS	B0878966
DEFINITION	B0878966 909 bp mRNA linear EST 16-AUG-2002
ACCESSION	AGNCSCURT.8183107 lupal.1 dorsal root ganglion Homo sapiens cDNA clone IMAGE:6184129 5' mRNA sequence.
VERSION	B0878966
KEYWORDS	B0878966.1 GI:22270974
	EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 909)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at:
<http://image.llnl.gov>
Plate: LIML3572 row: n column: 02
High quality sequence stop: 696.
Location/Qualifiers
1. 909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184129"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski dorsal root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1: NotI; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-GTGAGCTTGTGATGCGAGCGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 273 a 202 c 226 g 207 t 1 others

ORIGIN

Query Match 45 8%; Score 847.8; DB 13; Length 909;
Best Local Similarity 98.8%; Pred. No. 6.2e-151;
Matches 885; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

Qy 37 GTCCGCCGCTCAGAGCGCCCTATCAATTAATTAAAGAAACCACT-GGAAAAA 95
Db 1 GTCCGCCGCTCAGAGCGCCCTATCAATTAATTAAAGAAACCACTGGAAAAA 60

Qy 96 AAAATGAATTCCTTATCTTGCGATTTTGGGTGTTACCTTTATCCCTGTCTCT 155
Db 61 AAAATGAATTCCTTATCTTGCGATTTTGGGTGTTACCTTTATCCCTGTCTCT 120

Qy 156 GGGAAAGCTATATGCAAGATGCGATCTCTAAGAGACTTTGGAAGAAATAAAGAA 215
Db 121 GGGAAAGCTATATGCAAGATGCGATCTCTAAGAGACTTTGGAAGAAATAAAGAA 180

Qy 216 ATAGCAGCTGTGAGATGTTGCTAAACCAATCACTAGCTGTTATGTTAAAGCC 275
Db 181 ATAGCAGCTGTGAGATGTTGCTAAACCAATCACTAGCTGTTATGTTAAAGCC 240

Qy 276 CAGAAAGATCTTATAGCGATTTGCGACTTCTGTTGTAAGTGAACCAAGTGA 335
Db 241 CAGAAAGATCTTATAGCGATTTGCGACTTCTGTTGTAAGTGAACCAAGTGA 300

Qy 336 GGGTTCAGAACTTGAAGAAAGCCATCAATTAATGTAACAAAGCTGAGAGATGG 395
Db 301 GGGTTCAGAACTTGAAGAAAGCCATCAATTAATGTAACAAAGCTGAGAGATGG 360

Qy 396 CTGAGAAAGTTTCACTGAGAGCACTGAGAAATCCCACTGGAGAGGGAGAAATCA 455

Db 361 CTAGAGAAAGTTTCACTGAGAGCACTGAGAAATACCCCACTGGAGAGGGAGAAATCA 420
Qy 456 GCTGTATGCTGTGAGAGCCCAAGAAATTCATTAAGATGCCATCTTGCTTTGGCAGAGATT 515
Db 421 GCTGTATGCTGTGAGAGCCCAAGAAATTCATTAAGATGCCATCTTGCTTTGGCAGAGATT 480
Qy 516 GGGACTCTCTCAGAAAGGCTTTACAGCAGAGATTTCTGTGTGACCTCTTTGATGAATCG 575
Db 481 GGGACTCTCTCAGAAAGGCTTTACAGCAGAGATTTCTGTGTGACCTCTTTGATGAATCG 540
Qy 576 CAGAGAAAGGCTCTCAGAGAGAGAGAGAGATTTGTTATTAACAACCTTACATCAAC 635
Db 541 CAGAGAAAGGCTCTCAGAGAGAGAGAGAGATTTGTTATTAACAACCTTACATCAAC 600
Qy 636 TACTCAAGACGGTGCAATCCGAAAGCAGAGGGCGGTGGAAGCTGCCAAGGTGGGGCT 695
Db 601 TACTCAAGACGGTGCAATCCGAAAGCAGAGGGCGGTGGAAGCTGCCAAGGTGGGGCT 660
Qy 696 TTGGCATCTCTCATTCGATCCGTGGCTCCTCTCTCATCTACAGTCCCTCACACAGTATT 755
Db 661 TTGGCATCTCTCATTCGATCCGTGGCTCCTCTCTCATCTACAGTCCCTCACACAGTATT 720
Qy 756 CAGGAATACAGAGATGGCGTGGCCCAAAATTCACACAGCTGTATTACGGTGAAGATCA 815
Db 721 CAGGAATACAGAGATGGCGTGGCCCAAAATTCACACAGCTGTATTACGGTGAAGATCA 780
Qy 816 GAATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTTCACTGCTAAAGAT-GGG 874
Db 781 GAATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTTCACTGCTAAAGATGGG 840
Qy 875 GGCAGAACCTACCCAGATCTGATCTCTTCAACT-GTAGCAGAGATCACTGGG 929
Db 841 GGCAGAACCTACCCAGATCTGATCTCTTNCACACTGTGAGCAGAGATCACTGGG 896

RESULT 12
BX416895/c
LOCUS
DEFINITION
BX416895 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DB001YF04
3-PRIME, mRNA sequence.
VERSION
BX416895
KEYWORDS
BX416895.1 GI:30650311
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1003)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f. For more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DB001D02NP1&cluster=4663.f>. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/InvitrogenCorporation1600>
Faraday Avenue Genoscope sequence ID : CS0DB001D02NP1.
Location/Qualifiers
1. 1003
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB001YF04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched.

FEATURES
source

double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. The library was not normalized.

BASE COUNT	214 a	209 c	205 g	318 t	57 others
ORIGIN	.				

Query Match	45.7% ; Score 845.2 ; DB 13 ; Length 1003 ;
Best Local Similarity	88.1% ; Pred. No. 1,9e-150 ;
Matches	885 ; Conservative 36 ; Mismatches 80 ; Indels 3 ; Gaps 2 ;
QY 817	AAATGATGTCAAGAAATGGCTTCTCATGGGAATCAAAATTTGATTCAGTCAAGTAAGATGGGGG 876
Db 1003	AAATGATGTCAARAATKCTTCTCATGGATCAAAATTTGATTCAGTCAAGTAAGATGGGGG 944
QY 877	CAAAGACCTACCCAGATCTGATTTCTTCAACCTGTAGCAGATCATCTGGAGCAAT 936
Db 943	CAAAGACCTACCCAGAAACGTATTTCTTCAACCTGTAGCAGATCATCTGGAGCAAT 884
QY 937	ATCCAGAAACAGTTGTACTGTCTAGTGCATCTGCAAGCTGGGATGTTGGGACGGGTG 996
Db 883	ATCCAGAAACAGTTGTACTGTCTAGTGCATCTGCAAGCTGGGATGTTGGGACGGGTG 824
QY 997	CCATGATGATGCGCGGTGAGCCCTTTATATCATGGGAAGCACTTCATCTTATTAAGATC 1056
Db 823	CCATGATGATGCGCGGTGAGCCCTTTATATCATGGGAAGCACTTCATCTTATTAAGATC 764
QY 1057	TTGGGCGTCCGCCAAAGGGAATCTGCGGCTGTGCTTGGACATGCGCAAGAACAAAGTGT 1118
Db 763	TTGGGCGTCCGCCAAAGGGAATCTGCGGCGGTGCTTGGACATGCGCAAGAACAAAGTGT 704
QY 1117	GAGTTGGTGCCTTCAGATTAATTCAGTTACCAAGATTAATTTCCAACTACAGTCTGG 1176
Db 703	GAGTTGGTGCCTTACAGATTAATTCAG--NACCAAGAAATATTTCCAACTACAGTCTGG 645
QY 1177	TGATGAGTCTGACGCGAAGACCTTCTTACCACCTGGCTGCAATTCACCTGCGAGTAAA 1238
Db 644	TGAAGGAGTCTGACGCGAAGACCTTCTTACCACCTGGCTGCAATTCACCTGCGAGTAAA 585
QY 1237	AGGCGAGGGGCATCATGAGGAGGTATGAGCGCTGCGACGCCCTCAATTAATCTCAGG 1298
Db 584	AGGCGAGGGGCCTTATATGAGGAGGTATGAGCGCTGCGACGCCCTCAATTAATCTCAGG 525
QY 1297	TCCTGAGCCATGAGAA--GGGACGACATCAACTTTTGGATTCGACGCTGAGTGCCTGG 1354
Db 524	TCGAGGCAATGAGAAATGGGATCAGATCAATTTTGGAAACCAAGCTGGAGTGCMTGG 465
QY 1355	AGCCAGTCTACTGTGATGACTATTAACAGTATTTCTTCCATCATCTGCCACGAGACAC 1414
Db 464	ARCMAGTCAACTGTGAAAACCTAAAACAAMWATTTCCNNCAACCAAAACNCCACGGGACAC 405
QY 1415	CATGACGTGTCATGATCCCAAGCAGTGAATGTTGCTGCTGTTGGGCGCTTTGTTTC 1474
Db 404	CATGACGTGTCATGATCCCAAGCAGTGAATGTTGCTGCTGTTGGGCGCTTTGTTTC 345
QY 1475	TTATGTTGTCAGACATGAGAAATGCTGCTGATGCTTATGAAAACAGTAAAGAAAGAA 1534
Db 344	TNAGGAAGMACAACAAGAAAGAAAGAACACCTAGTGTCAAGAAACAGTAAAGAAAGAA 285
QY 1535	CGTTTTCATGCTTCTGCGCCAGGAATCTGGGTCTGCAACTTTGGAAAATCCTCTTTCACA 1594
Db 284	SGTTTTCATGCTTCTGCGCCAGGAATCTGGGTCTGCAACTTTGGAAAATCCTCTTTCACA 225
QY 1595	TAAACAATTTCAATCCAAATTCATCTTCAAGCAACTCTAATTTCAATGCTTTCGTTAATAT 165
Db 224	TAAACAATTTCAATCCAAATTCATCTTCAAGCAACTCTAATTTCAATGCTTTCGTTAATAT 165
QY 1655	CTTTCTTGATACCTTCCAAATTTCTGATTTCTAGAAAAAGGAATCATTTCTCCCTCCTC 1714
Db 164	CTTTCTTGATACCTTCCAAATTTCTGATTTCTAGAAAAAGGAATCATTTCTCCCTCCTC 105
QY 1715	CCACCACTAATAATCAATATGTAAGGATTTACGTGGGGGCAATTTCTTATATCACT 1774
Db 104	CTTCACTTAATAATCAATATGTAAGGATTTACGTGGGGGCAATTTCTTATATCACT 45

Db	Accession	Version	KeyWords	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	BX416896	1	5-PRIME, mRNA sequence.	EST.	Homo sapiens (human)	1 (bases 1 to 1012)	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DB001DC02QP1&cluster=4663.f. Contact : Feng Liang Email : fliang@life.techn.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue genoscope sequence ID : CS0DB001DC02P1.
FEATURES	Source	Location/Qualifiers								
	1..1012	/organism="Homo sapiens"								
	/mol_type="mRNA"									
	/db_xref="taxon:9606"									
	/clone="CS0DB001YF04"									
	/cissue_type="PLACENTA"									
	/clone_id="Homo sapiens PLACENTA"									
	/note="Vector: pCMVSPORT 6, 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."									
BASE COUNT	281 a	240 c	275 g	215 t	1 others					
ORIGIN										
Query Match	45.4%	Score 840.8	DB 13	Length 1012						
Best Local Similarity	99.6%	Pred. No. 1.3e-149								
Matches 853	Conservative 0	Mismatches 2	Indels 1	Gaps 1						
Db	1	GGCTAGGCGCGAGCTTAGTCTCTGGAGCGCGCTCGTCGCGCGCGCTCAGAGCGGCCCTA	60							
Db	157	GGCTAGGCGCGAGCTTAGTCTCTGGAGCGCGCTCGTCGCGCGCGCTCAGAGCGGCCCTA	216							
Db	61	TCAGATTATCTTAAACAGAAAACCACTGGAGAAAAAATGAATTCCTTATCTTCGCAT	120							
Db	217	TCAGATTATCTTAAACAGAAAACCACTGGAGAAAAAATGAATTCCTTATCTTCGCAT	276							
Db	121	TTTTCGATGGGTGCTTACCTTTATCCCTGTGCTGTGGAAAGCATATGCAAGAAATGGCA	180							
Db	277	TTTTCGATGGGTGCTTACCTTTATCCCTGTGCTGTGGAAAGCATATGCAAGAAATGGCA	336							
Db	181	TTCTTAAGAGAGCTTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGAGATGTGCTA	240							
Db	337	TTCTTAAGAGAGCTTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGAGATGTGCTA	396							
Db	241	AAGCAATCATCACTGCTGTTTATAGTAAAGCCGAAACAGATCTCTAGAGGATTTGG	300							
Db	397	AAGCAATCATCACTGCTGTTTATAGTAAAGCCGAAACAGATCTCTAGAGGATTTGG	456							

QY 301 CACTTCTGGTGTATCTGTGGACCCAGACTAGTGGCTCCAAGAACTAGAAAAAGCCA 360
 DB 457 CACTTCTGGTGTATCTGTGGACCCAGACTAGTGGCTCCAAGAACTAGAAAAAGCCA 516
 QY 361 TCCTCAATTAATGATCAAAACCTGCGACGAAAGATGGCTGAGAAAGTTCACTGGAGCCAG 420
 DB 517 TCCTCAATTAATGATCAAAACCTGCGACGAAAGATGGCTGAGAAAGTTCACTGGAGCCAG 576
 QY 421 TGAGAAATACCCCACTGGGAGAGGGGAGAAATACAGCTGTGATGCTGGAGCCAAATTC 480
 DB 577 TGAGAAATACCCCACTGGGAGAGGGGAGAAATACAGCTGTGATGCTGGAGCCAAATTC 636
 QY 481 ATAGATAGCCATCTCTGGCTCTTGGACAGCATTTGGACCTCTCCAGAGGCAATTACAG 540
 DB 637 ATAGATAGCCATCTCTGGCTCTTGGACAGCATTTGGACCTCTCCAGAGGCAATTACAG 696
 QY 541 CAGAAATTTGGTGTGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGAGAG 600
 DB 697 CAGAAATTTGGTGTGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGAGAG 756
 QY 601 GGAAGATTTGTTTATATACCAACCTTATCATCACTCAAGAGAGGCTGCAATACCGAA 660
 DB 757 GGAAGATTTGTTTATATACCAACCTTATCATCACTCAAGAGAGGCTGCAATACCGAA 816
 QY 661 CGCAGGGGGCGGTGGAAGCTGCGCAAGG-TGGGGGCTTTGGCATCTCTCATTCGATCCG 719
 DB 817 CGCAGGGGGCGGTGGAAGCTGCGCAAGG-TGGGGGCTTTGGCATCTCTCATTCGATCCG 876
 QY 720 GCTCTCTTCCATCTACAGTCTCTCAACAGATTTACAGAAATACAGAGTGGCTGCC 779
 DB 877 GCTCTCTTCCATCTACAGTCTCTCAACAGATTTACAGAAATACAGAGTGGCTGCC 936
 QY 780 AAAATTTCAACAGCCGTATTAAGGTGGAAGATGAGAAATGATGCAAGATGGCTCT 839
 DB 937 AAGATTTCAACAGCCGTATTAAGGTGGAAGATGAGAAATGATGCAAGATGGCTCT 996
 QY 840 CATGGATCAAAATTTG 855
 DB 997 CATGGATCAAAATTTG 1012

RESULT 14
 LOCUS BX335995 919 bp mRNA linear EST 02-MAY-2003
 DEFINITION BX335995 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1022YJ16 5-PRIME, mRNA sequence.
 ACCESSION BX335995
 VERSION BX335995.1 GI:30339459
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 919)
 JOURNAL li.W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 invitrogen. This sequence belongs to sequence cluster 4663.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D1022DE08QPlcluster=4663.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1022DE08QPl.

FEATURES
 SOURCE

1..919
 /organism="Homo sapiens"

/mol type="mRNA"
 /db xref="taxon:9606"
 /clone="CS0D1022YJ16"
 /rissue type="PLACENTA COT 25-NORMALIZED"
 /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-0.150 (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 273 a 206 c 226 g 208 t 6 others
 ORIGIN

Query Match 45.3%; Score 838.6; DB 13; Length 919;
 Best Local Similarity 96.9%; Pred. No. 3.4e-149;
 Matches 861; Conservative 23; Mismatches 23; Indels 1; Gaps 1;
 QY 1 GCTAGGCCGCGACCTTATGTCCTGGAGCCGCTCCGTCGCGCTCAGAGCCGCTTA 60
 DB 28 GCGTGTATACCCGCTCCGGAATTCGCGGATCGCTCCGTCGCGCTCAGAGCCGCTTA 87
 QY 61 TCAGATTAATCTTAACAAAGAAACCAACTGAAAGAAAGAAATTCCTTATCTTCCAT 120
 DB 88 TCAGATTAATCTTAACAAAGAAACCAACTGAAAGAAAGAAATTCCTTATCTTCCAT 147
 QY 121 TTTTCGGTGTGTACCTTTTATTCCTGCTCTGCGAAAGCTATATGCAAGATGCA 180
 DB 148 TTTTCGGTGTGTACCTTTTATTCCTGCTCTGCGAAAGCTATATGCAAGATGCA 207
 QY 181 TCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGACCGCTGTGAGATGTTGCTA 240
 DB 208 TCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGACCGCTGTGAGATGTTGCTA 267
 QY 241 AAGCAATCATCACTGCTGTTATGTTAAGCCGAAACAGATCTCTATAGAGATTTG 300
 DB 268 AAGCAATCATCACTGCTGTTATGTTAAGCCGAAACAGATCTCTATAGAGATTTG 327
 QY 301 CACTTCTGTGATATCTGTGACCCAGACTAGTGGCTCCAGAACTAGAAAAAGCCA 360
 DB 328 CACTTCTGTGATATCTGTGACCCAGACTAGTGGCTCCAGAACTAGAAAAAGCCA 387
 QY 361 TCCTCAATTAATGATCAAAACCTGCGACGAAAGATGGCTGAGAAAGTTCACTGGAGCCAG 420
 DB 388 TCCTCAATTAATGATCAAAACCTGCGACGAAAGATGGCTGAGAAAGTTCACTGGAGCCAG 447
 QY 421 TGAGAAATACCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATGCTGGAGCCAAATTC 480
 DB 448 TGAGAAATACCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATGCTGGAGCCAAATTC 507
 QY 481 ATAGATAGCCATCTCTGGCTCTTGGACAGCATTTGGACCTCTCCAGAGGCAATTACAG 540
 DB 508 ATAGATAGCCATCTCTGGCTCTTGGACAGCATTTGGACCTCTCCAGAGGCAATTACAG 567
 QY 541 CAGAAATTTGGTGTGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGAGAG 600
 DB 568 CAGAAATTTGGTGTGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGAGAG 627
 QY 601 GGAAGATTTGTTTATATACCAACCTTATCATCACTCAAGAGAGGCTGCAATACCGAA 660
 DB 628 GGAAGATTTGTTTATATACCAACCTTATCATCACTCAAGAGAGGCTGCAATACCGAA 687
 QY 661 CGCAGGGGGCGGTGGAAGCTGCGCAAGG-TGGGGGCTTTGGCATCTCTCATTCGATCCG 720
 DB 688 CGCAGGGGGCGGTGGAAGCTGCGCAAGG-TGGGGGCTTTGGCATCTCTCATTCGATCCG 747
 QY 721 CCTCTCTTCCATCTACAGTCTCTCAACAGATTTCAAGAAATACAGAGATGGCTGCCA 780
 DB 748 CCTCTCTTCCATCTACAGTCTCTCAACAGATTTCAAGAAATACAGAGATGGCTGCCA 807
 QY 781 AAATTTCAACAGCTGTATTAAGGTGGAAGATGAGAAATGATGCAAGATGGCTCTC 840
 DB 808 AGATTTCAACAGCTGTATTAAGGTGGAAGATGAGAAATGATGCAAGATGGCTCTC 867
 QY 841 AT-GGATCAAAATTTGTCATTCAGCTAAGATGGGGCAAGACCTTACC 888

Db 868 ATGGGGATCAAAATGTCATTACACTAAARATGGGGGACARACTTACC 916
 RESULT 15
 LOCUS BX335994/c 929 bp mRNA linear EST 02-MAY-2003
 DEFINITION BX335994 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
 clone CSOD1022YU16 3-PRIME, mRNA sequence.
 ACCESSION BX335994
 VERSION BX335994.1 GI:30337474
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 929)
 Full-length cDNA libraries and normalization
 Unpublished
 CONTACT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4663.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSOD1022DE08NP1&cluster=4663.f. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSOD1022DE08NP1.
 FEATURES
 source
 1..929
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1022YU16"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 268 a 200 c 198 g 234 t 23 others
 ORIGIN
 Query Match 44.9%; Score 831.6; DB 13; Length 929;
 Best Local Similarity 96.3%; Pred. No. 7.3e-148;
 Matches 868; Conservative 12; Mismatches 18; Indels 3; Gaps 3;
 QY 900 TCCTTCAACACTGACGAGATCACTGGAGCAATATTCAG-AACAGGTTGATGCT 958
 Db 929 TCCTTCAACACTGACGAGATCACTGGAGCAATATTCAG-AACAGGTTGATGCT 870
 QY 959 CAGTGCATCTGACAGCTGGAGTGTGGGC-AGGGTGCATGATGATGCGGTGAG 1017
 Db 869 CAGTGCATCTGACAGCTGGAGTGTGGGC-AGGGTGCATGATGATGCGGTGAG 810
 QY 1018 CTTTATATCATGGAAGACTCTCACTTATTAAGATCTGGGCTGGCTCC-AAAGAG 1076
 Db 809 CTTTATATCATGGAAGACTCTCACTTATTAAGATCTGGGCTGGCTCC-AAAGAG 750
 QY 1077 ACTCTGGGGCTGTGCTCTGAGTGCAGAGAACAGGTGAGTGTGCTTCAGTAT 1136
 Db 749 ACTCTGGGGCTGTGCTCTGAGTGCAGAGAACAGGTGAGTGTGCTTCAGTAT 690
 QY 1137 TATCAGTTACACAGGTAAATTTCCATACAGTCTGTGATGAGATCTGACGCGAGA 1196
 Db 689 TATCAGTTACACAGGTAAATTTCCATACAGTCTGTGATGAGATCTGACGCGAGA 630
 QY 1197 ACCTTCTTACCACCTGGGCTCAATTCACTGGCAGTGAAGCCAGGGCCATCATGAG 1256

Db 629 ACCTTCTTACCACCTGGGCTCAATTCACTGGCAGTGAAGAAAGCCAGGGCCATCATGAG 570
 QY 1257 GAGGTATGAGCCCTGCTGACAGCCCTCAATATCACTGAGGTCCTGAGCCATGAGAGAGG 1316
 Db 569 GAGGTATGAGCCCTGCTGACAGCCCTCAATATCACTGAGGTCCTGAGCCATGAGAGAGG 510
 QY 1317 ACAGACATCAACTTTTGGATTCAGAGCTGAGTGCCTGAGCCAGTCTACTGTATGACTTA 1376
 Db 509 ACAGACTTCACTTTTGGATTCAGAGCTGAGTGCCTGAGCCAGTCTACTGTATGACTTA 450
 QY 1377 TACAAGTATTTCTTCTTCATCACTCCACGAGACACCAATGACTGATGATCCAAAG 1436
 Db 449 TACAAGTATTTCTTCTTCATCACTCCACGAGACACCAATGACTGATGATCCAAAG 390
 QY 1437 CAGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1496
 Db 389 CAGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330
 QY 1497 GAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1556
 Db 329 GAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 270
 QY 1557 AATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1616
 Db 269 AATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 210
 QY 1617 TTCAAAGCACAACCTTATTTTCACTGCTTCTGTTATATCTTCTGTAATCTTCCAAAT 1676
 Db 209 TTCAAAGCACAACCTTATTTTCACTGCTTCTGTTATATCTTCTGTAATCTTCCAAAT 150
 QY 1677 CTCTGATTTAGAAAAAGAAATCAATCTCCCTCCCTCCGACCAATAGAAATCAATAT 1736
 Db 149 CTCTGATTTAGAAAAAGAAATCAATCTCCCTCCCTCCGACCAATAGAAATCAATAT 90
 QY 1737 GGTAGGATTAACAGTGGGGCAATTTCTTATATCACTCTTAAACATTTGTTCACTT 1796
 Db 89 GGTAGGATTAACAGTGGGGCAATTTCTTATATCACTCTTAAACATTTGTTTCTT 30
 QY 1797 T 1797
 Db 29 T 29

Search completed: December 22, 2003, 23:11:23
 Job time : 3636.25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:59:40 ; Search time 122.256 Seconds
(without alignments)
6682.695 Million cell updates/sec

Title: US-09-745-763-35
Perfect score: 1851
Sequence: 1 GCGTAGGCGCCGAGCTTAGT.....CTGAAAAAAAAAAAAAAAAAAAA 1851

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1846.2	99.7	1863	US-09-482-273-28	Sequence 28, Appl
2	1098.4	59.3	1134	US-09-482-273-95	Sequence 95, Appl
3	49.2	2.7	7218	US-08-232-463-14	Sequence 14, Appl
4	40	2.2	646	US-08-487-001A-50	Sequence 50, Appl
5	40	2.2	646	US-08-630-822A-50	Sequence 50, Appl
6	40	2.2	646	US-09-005-069-50	Sequence 50, Appl
7	40	2.2	7218	US-08-232-463-14	Sequence 14, Appl
8	38	2.1	2049	US-08-268-797-1	Sequence 1, Appl
9	38	2.1	2049	PCT-US95-08414-1	Sequence 1, Appl
10	38	2.1	2691	US-07-878-960-1	Sequence 1, Appl
11	37.8	2.0	29793	US-09-302-812-38	Sequence 38, Appl
12	37.8	2.0	29793	US-09-511-477-38	Sequence 38, Appl
13	37.8	2.0	29793	US-09-511-507-38	Sequence 38, Appl
14	36.8	2.0	2393	US-09-016-000-11	Sequence 11, Appl
15	36	1.9	2854	US-09-996-243-66	Sequence 66, Appl
16	36	1.9	43360	US-09-453-702B-206	Sequence 206, App
17	36	1.9	45325	US-09-453-702B-261	Sequence 261, App
18	35.8	1.9	1443	US-09-904-615-21	Sequence 21, Appl
19	35.6	1.9	928	US-09-152-060-46	Sequence 46, Appl
20	35.6	1.9	1844	US-09-123-912-88	Sequence 88, Appl
21	35.6	1.9	1844	US-09-643-597-88	Sequence 88, Appl
22	35.6	1.9	1844	US-09-480-884A-88	Sequence 88, Appl
23	35.6	1.9	1844	US-09-543-615A-88	Sequence 88, Appl
24	35.6	1.9	1844	US-09-606-421B-88	Sequence 88, Appl
25	35.4	1.9	9048	US-08-973-273-4	Sequence 4, Appl
26	35.2	1.9	2373	US-08-980-080-1	Sequence 1, Appl
27	35.2	1.9	3003	US-08-434-730-15	Sequence 15, Appl

C	28	34.8	1.9	5340	4	US-09-627-122-21	Sequence 21, Appl
	29	34.6	1.9	2242	3	US-09-400-742-1	Sequence 1, Appl
	30	34.6	1.9	2242	3	US-08-618-651A-1	Sequence 1, Appl
	31	34.6	1.9	2242	4	US-09-215-252-1	Sequence 1, Appl
C	32	34.6	1.9	4086	4	US-09-220-132-56	Sequence 56, Appl
C	33	34.6	1.9	4334	4	US-09-620-312D-88	Sequence 88, Appl
C	34	34.4	1.9	738	5	US-08-738-462-1	Sequence 1, Appl
C	35	34.4	1.9	738	5	PCT-US94-07587-1	Sequence 1, Appl
	36	34	1.8	2106	4	US-09-252-991A-7477	Sequence 7477, Ap
	37	34	1.8	2112	4	US-09-252-991A-7334	Sequence 7334, Ap
C	38	34	1.8	2187	4	US-09-252-991A-7167	Sequence 7167, Ap
C	39	34	1.8	2757	4	US-09-482-273-74	Sequence 74, Appl
C	40	34	1.8	5035	2	US-08-616-392C-3	Sequence 3, Appl
	41	34	1.8	90541	4	US-09-759-359A-3	Sequence 1, Appl
C	42	34	1.8	1664976	4	US-08-916-421B-1	Sequence 1, Appl
C	43	33.8	1.8	2636	2	US-08-820-170A-21	Sequence 21, Appl
C	44	33.8	1.8	2636	3	US-09-055-699-21	Sequence 21, Appl
C	45	33.8	1.8	2636	3	US-09-273-565-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-482-273-28
Sequence 28, Application US/09482273
Patent No. 6534631
GENERAL INFORMATION:
APPLICANT: Roese et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
CURRENT FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 28
LENGTH: 1863
TYPE: DNA
ORGANISM: Homo sapiens
US-09-482-273-28
Query Match 99.7%; Score 1846.2; DB 4; Length 1863;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1848; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GCGTAGGCGCCGAGCTTAGTCTGGAGCGCGCTCGTCCGCCGCTGACAGCCGCCCTTA 60
DB 1 GACTAGGCGCGAGCTTAGTCTGGAGCGCGCTCGTCCGCCGCTGACAGCCGCCCTTA 60
QY 61 TCAGTATCTTTAACAAGAAACCACTGGGAAAAAATGAATTCCTATCTTCGCAT 120
DB 61 TCAGTATCTTTAACAAGAAACCACTGGGAAAAAATGAATTCCTATCTTCGCAT 120
QY 121 TTTTGGTGGGTTCACCTTTTATCCCTGCTGGGAAAGCTATGCAAGATGGCA 180
DB 121 TTTTGGTGGGTTCACCTTTTATCCCTGCTGGGAAAGCTATGCAAGATGGCA 180
QY 181 TCTTAAGAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATTTGCTA 240
DB 181 TCTTAAGAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATTTGCTA 240
QY 241 AAGCAATCATCAACTGCTGTTTATGTGTAAGCCCGAAGACAGTCTTATGAGGATTCG 300
DB 241 AAGCAATCATCAACTGCTGTTTATGTGTAAGCCCGAAGACAGTCTTATGAGGATTCG 300

```

OY 301 CACTTCGTGTGATACCTGTGGAGCCAGACGTAGTGCTCCAGAAAGCTAGAAAAAGCA 360
Db 301 CACTTCGTGTGATACCTGTGGAGCCAGACGTAGTGCTCCAGAAAGCTAGAAAAAGCA 360
OY 361 TCCAATATATGTATACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAGCCAG 420
Db 361 TCCAATATATGTATACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAGCCAG 420
OY 421 TGAAGATCCCCACCTGGGAGAGGGAGAAAGATCGCTGTATGTGTGAGAGCCAAATTC 480
Db 421 TGAAGATCCCCACCTGGGAGAGGGAGAAAGATCGCTGTATGTGTGAGAGCCAAATTC 480
OY 481 ATAGATATGACCATCTGGGCTCTTGGCAGCAGCATGGAGCTCTCCAGAGGCAATTACAG 540
Db 481 ATAGATATGACCATCTGGGCTCTTGGCAGCAGCATGGAGCTCTCCAGAGGCAATTACAG 540
OY 541 CAGAAATTCGTGGGTGACCTCTTTCGATGAACCTGCAAGAAAGGCTCCAGAGCAGAG 600
Db 541 CAGAAATTCGTGGGTGACCTCTTTCGATGAACCTGCAAGAAAGGCTCCAGAGCAGAG 600
OY 601 GGAAGATGTGTGTTATATACCAACCTTACATCACTCAAGAGCGGTGCAATCCGAA 660
Db 601 GGAAGATGTGTGTTATATACCAACCTTACATCACTCAAGAGCGGTGCAATCCGAA 660
OY 661 CGCAGGGGGCGGTGAGAGCTGCAGAGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 720
Db 661 CGCAGGGGGCGGTGAGAGCTGCAGAGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 720
OY 721 CCTCCCTTCATCTACAGTCTCTCAACAAGTATTCAGAAATCCAGAGATGGCGTCCCA 780
Db 721 CCTCCCTTCATCTACAGTCTCTCTCAACAAGTATTCAGAAATCCAGAGATGGCGTCCCA 780
OY 781 AAATTCCAACAGCCCTGTATATAGGTGAGAAAGATGCAAAATGATGTCAGAAATGCTTCTC 840
Db 781 AGATTCCAACAGCCCTGTATATAGGTGAGAAAGATGCAAAATGATGTCAGAAATGCTTCTC 840
OY 841 ATGGGATCAAAATGTCTCATTCAGTAAAGATGGGGCAAAAGCTAACCAATCTAGATT 900
Db 841 ATGGGATCAAAATGTCTCATTCAGTAAAGATGGGGCAAAAGCTAACCAATCTAGATT 900
OY 901 CCTTCACACGTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAAGGTTGTACTGTGCA 960
Db 901 CCTTCACACGTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAAGGTTGTACTGTGCA 960
OY 961 GTGACATCTGTGACAGCTGGGATGTTGGGAGAGGTCATGATATGCGGTGAGGCT 1020
Db 961 GTGACATCTGTGACAGCTGGGATGTTGGGAGAGGTCATGATATGCGGTGAGGCT 1020
OY 1021 TTAATTCATGGGAGACCTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGGACTC 1080
Db 1021 TTAATTCATGGGAGACCTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGGACTC 1080
OY 1081 TGGGGCTGTGTCTGTGACCTGCAAGAAAGCAAGGTGAGGTGGTCTTCCAGATATATC 1140
Db 1081 TGGGGCTGTGTCTGTGACCTGCAAGAAAGCAAGGTGAGGTGGTCTTCCAGATATATC 1140
OY 1141 AGTTACACAAGGTAAATTTTCCAACTACGTCTGTATGAGTCTTACCGAGAACT 1200
Db 1141 AGTTACACAAGGTAAATTTTCCAACTACGTCTGTATGAGTCTTACCGAGAACT 1200
OY 1201 TCTTACCACTGGGCTGCAATTCATCTGGCAGTGAAGAGGCGGCATCATGAGGAGG 1260
Db 1201 TCTTACCACTGGGCTGCAATTCATCTGGCAGTGAAGAGGCGGCATCATGAGGAGG 1260
OY 1261 TTATAGGCTGTGAGCCCTCAATATCACTCAGAGTCTTGAAGCAGAGAGGAGAG 1320
Db 1261 TTATAGGCTGTGAGCCCTCAATATCACTCAGAGTCTTGAAGCAGAGAGGAGAGAG 1320
OY 1321 ACATTAACCTTTTGTGATCCAGAGTGGAGTCTGAGAGGCAAGTTACTTGAATCTTATCA 1380
Db 1321 ACATTAACCTTTTGTGATCCAGAGTGGAGTCTGAGAGGCAAGTTACTTGAATCTTATCA 1380
OY 1381 AGTATTTCTTCTTCATCACTCCAGAGGAGACACATGATCTGTATGATCCAAAGCAG 1440

```

```

Db 1381 AGTATTTCTTCTTCATCACTCCAGAGGAGACACATGATGTCATGATCCAAAGCAG 1440
OY 1441 TGAATGTGTGCTGCTGCTGTTTGGGCTGTGTTCTTATATGTTGTGACAGATGGAAGAA 1500
Db 1441 TGAATGTGTGCTGCTGCTGTTTGGGCTGTGTTCTTATATGTTGTGACAGATGGAAGAA 1500
OY 1501 TGTGCTAGAGTCTTGAAGAACAGTGAAGAAAGAAAGCTTTTCATGCTCTGGCCAGAAATC 1560
Db 1501 TGTGCTAGAGTCTTGAAGAACAGTGAAGAAAGAAAGCTTTTCATGCTCTGGCCAGAAATC 1560
OY 1561 CTGGGTGTGCACTTTGAAAACCTCTTTCATTAACATTAATTCATCAATTCCTTCA 1620
Db 1561 CTGGGTGTGCACTTTGAAAACCTCTTTCATTAACATTAATTCATCAATTCCTTCA 1620
OY 1621 AAGACAACCTATTTACAGCTTCTGTTATATATCTTCTGATACCTTCCAAATTCCT 1680
Db 1621 AAGACAACCTATTTATATGCTGTTATATATCTTCTGATACCTTCCAAATTCCT 1680
OY 1681 GATTCCTGAAAAGAGATCAATTCCTCCCTCCCTCCACACATGGAATCAATATGATA 1740
Db 1681 GATTCCTGAAAAGAGATCAATTCCTCCCTCCCTCCACACATGGAATCAATATGATA 1740
OY 1741 GGAATTAAGTGGGGGCAATTCCTTATATCACTCTTAAACATTTGTTCACTTAAA 1800
Db 1741 GGAATTAAGTGGGGGCAATTCCTTATATCACTCTTAAACATTTGTTCACTTAAA 1800
OY 1801 AGTAAACACTTAATTAATTTTGGAGAGATCTTGAAAAAAGAAAAA 1851
Db 1801 AGTAAACACTTAATTAATTTTGGAGAGATCTTGAAAAAAGAAAAA 1851

RESULT 2
US-09-482-273-95
; Sequence 95, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OR INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; EARLIER FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 95
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-482-273-95

Query Match 59.3%; Score 1098.4; DB 4; Length 1134;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1121; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

OY 729 TCCATCTACAGTCTCTCAACAAGTATTCAGAAATACAGAGATGGCGCCAAATTC 788
Db 1 TCCATCTACAGTCTCTCAACAAGTATTCAGAAATACAGAGATGGCGCCAAATTC 60
OY 789 ACAGCCTGTATTTAGAGTGAAGATGCAAGAAATGATGTCAGAAATGAGCTTCTCATGGATC 848
Db 61 ACAGCCTGTATTTAGAGTGAAGATGCAAGAAATGATGTCAGAAATGAGCTTCTCATGGATC 120
OY 849 AAAATTTGTCATTCAGCTTAAGATGGGGCAAAAGACTTACCCAGATCTGATTCCTTCAAC 908
Db 121 AAAATTTGTCATTCAGCTTAAGATGGGGCAAAAGACTTACCCAGATCTGATTCCTTCAAC 180

```


Qy 567 GATGAAGTCAGAGAGGCTCTCA 590
Db 1051 GACCTGACGACCAAGCTCGGAATTA 1028

RESULT 4

US-08-487-001A-50

Sequence 50, Application US/08487001A

Patent No. 5795862

GENERAL INFORMATION:

APPLICANT: FRANK, GLENN R.

APPLICANT: HUNTER, SHIRLEY WU

APPLICANT: WALLENFELS, LYNDA

TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA

TITLE OF INVENTION: PROTEINS AND APPARATUS TO COLLECT SUCH PROTEINS

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: U.S.A.

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,001A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Verser, Carol Talkington

REGISTRATION NUMBER: 37,459

REFERENCE/DOCKET NUMBER: 2618-17-C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 646 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 3..519

US-08-487-001A-50

Query Match 2.2%; Score 40; DB 1; Length 646;

Best Local Similarity 63.5%; Pred. No. 0.042;

Matches 61; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1756 GCATTTCTTATATACCTCTTAACCACTGTTCCACTTAAGTAACCTTAATA 1815

Db 551 GCTATTTTCTGTAACCATATTAAGCTATTTTAATACCTTGACGATATATACATAATA 610

Qy 1816 AATTTTGGAGATCTCTGAAAAA 1851

Db 611 AATTGCTACATTGCTCTTAACCACTGTTCCACTTAAGTAACCTTAATA 610

Qy 1816 AATTTTGGAGATCTCTGAAAAA 1851

Db 611 AATTGCTACATTGCTCTTAACCACTGTTCCACTTAAGTAACCTTAATA 610

RESULT 5

US-08-630-822A-50

Sequence 50, Application US/08630822A

Patent No. 5840695

GENERAL INFORMATION:

APPLICANT: FRANK, GLENN R.

APPLICANT: HUNTER, SHIRLEY WU

APPLICANT: WALLENFELS, LYNDA

TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross P.C.

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: U.S.A.

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,822A

FILING DATE: 11-APR-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CONNELL, GARY J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-17-C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 646 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 3..519

US-08-630-822A-50

Query Match 2.2%; Score 40; DB 2; Length 646;

Best Local Similarity 63.5%; Pred. No. 0.042;

Matches 61; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1756 GCATTTCTTATATACCTCTTAACCACTGTTCCACTTAAGTAACCTTAATA 1815

Db 551 GCTATTTTCTGTAACCATATTAAGCTATTTTAATACCTTGACGATATATACATAATA 610

Qy 1816 AATTTTGGAGATCTCTGAAAAA 1851

Db 611 AATTGCTACATTGCTCTTAACCACTGTTCCACTTAAGTAACCTTAATA 610

Qy 1816 AATTTTGGAGATCTCTGAAAAA 1851

Db 611 AATTGCTACATTGCTCTTAACCACTGTTCCACTTAAGTAACCTTAATA 610

RESULT 6

US-09-005-069-50

Sequence 50, Application US/09005069

Patent No. 5932470

GENERAL INFORMATION:

APPLICANT: FRANK, GLENN R.

APPLICANT: HUNTER, SHIRLEY WU

APPLICANT: WALLENFELS, LYNDA

TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS

TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS

NUMBER OF SEQUENCES: 107

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross P.C.

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: U.S.A.

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

RESULT 7
 US-08-232-463-14
 Sequence 14, Application US/08232463
 Patent No. 5670367
 GENERAL INFORMATION:
 APPLICANT: DORNER, F.
 APPLICANT: SCHEIFLINGER, F.
 APPLICANT: FALKNER, F. G.
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232,463
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935,313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.

?
?
? RESULT 8
? US-08-268-797-1/c
? Sequence 1, Application US/08268797
? Patent No. 5599788
? GENERAL INFORMATION:
? APPLICANT: Purchio, Anthony F.
? APPLICANT: Lebaron, Richard
? TITLE OF INVENTION: Factor to Grow Tissue Ex Vivo
? NUMBER OF SEQUENCES: 1
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Knobbe, Martens, Olson and Bear
? STREET: 620 Newport Center Drive, Sixteenth Floor
? CITY: Newport Beach
? STATE: CA
? COUNTRY: USA
? ZIP: 92660
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/268,797
? FILING DATE:
?

1 APPLICANT: Purchio, Anthony F.
2 APPLICANT: Skonier, John
3 APPLICANT: Neuhauer, Michael G.
4 TITLE OF INVENTION: TGF-BETA INDUCED GENE AND PROTEIN
5 NUMBER OF SEQUENCES: 2
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: Bristol-Myers Squibb Company
8 STREET: 3005 First Avenue
9 CITY: Seattle
10 STATE: Washington
11 COUNTRY: USA
12 ZIP: 98121
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: Patent Release #1.0, Version #1.25
19
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/07/878,960
22 FILING DATE: 05-MAY-1992
23 CLASSIFICATION: 530
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US 07/833,835
26 FILING DATE: 05-FEB-1992
27
28 NAME: 530
29
30 ATTORNEY/AGENT INFORMATION:
31 NAME: Sorrentino, Joseph M.
32 REGISTRATION NUMBER: 32,598
33 REFERENCE/DOCKET NUMBER: ON0092-
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: 206/728-4800
36 TELEFAX: 206/727-3601
37
38 INFORMATION FOR SEQ ID NO: 1:
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 2691 base pairs

Patent No. 6478825

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerlitsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gunney, Austin L.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: P2730P1C13

CURRENT FILING DATE: 2001-11-14

PRIOR APPLICATION NUMBER: US/09/996,243

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/065186

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/075945

PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/084600

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/087106

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087609

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087759

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087827

PRIOR FILING DATE: 1998-06-03

PRIOR APPLICATION NUMBER: 60/088021

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088025

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088026

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088028

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088029

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088030

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088033

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 1.9%; Score 36; DB 4; Length 2854;
Best Local Similarity 67.1%; Pred. No. 2;
Matches 51; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1776 TTAATAACATGTTTCACCTTAATAAGTAACCTTAATAATTTTGGAGATCTCGA 1835
DB 2778 TTAATGCTTACCTCCCTTTAAATTAATGATTAATAATGCTCTTGAATAAAAAA 2837
QY 1836 AAAAAAAAAAAAAA 1851
DB 2838 AAAAAAAAAAAAAA 2853

Search completed: December 22, 2003, 23:17:17
Job time : 129.256 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 14:49:05 ; Search time 579.108 Seconds
(without alignments)
10650.710 Million cell updates/sec

Title: US-09-745-763-35
Perfect score: 1851
Sequence: 1 GCGTAGGCGCCGAGCTTAGT.....CTGAAAAAAAAAAAAAAAAA 1851

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2211978 seqs, 1666101734 residues
Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/2/pubpna/PCRUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 - 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
 - 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
 - 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 - 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1851	100.0	1851	9	US-09-745-763-35
2	1846.2	99.7	1863	11	US-09-984-271-28
3	1145	61.9	1778	10	US-09-917-800A-505
4	1098.4	59.3	1134	11	US-09-984-271-95
5	261.8	14.1	357	10	US-09-833-381-1929
6	247.4	13.4	427	10	US-09-833-381-1930
7	233	12.6	234	10	US-09-880-107-3881
8	135	7.3	546	13	US-10-027-632-192834
9	135	7.3	546	14	US-10-027-632-192834
10	91	4.9	134	10	US-09-783-590-12141
11	60.2	3.3	65	13	US-09-908-975-25931
12	60	3.2	60	13	US-09-908-975-8960
13	45.4	2.5	40862	13	US-10-311-455-2046
14	42	2.3	2200	15	US-10-059-585-5
15	41.8	2.3	13202	13	US-10-311-455-1458

C	16	41.2	2.2	328	13	US-10-076-747-72
	17	41.2	2.2	664	13	US-09-814-353-6384
	18	41.2	2.2	664	13	US-09-814-353-12661
C	19	41.2	2.2	7384	13	US-10-311-455-720
	20	41	2.2	65	13	US-09-908-975-1202
C	21	41	2.2	445	13	US-09-814-353-17408
C	22	40.4	2.2	3673778	13	US-10-312-841-2
C	23	40.2	2.2	397	13	US-09-814-353-17565
C	24	40.2	2.2	8946	13	US-10-311-455-884
	25	40	2.2	646	8	US-08-809-423-50
	26	40	2.2	646	13	US-10-271-344-50
C	27	39.6	2.1	881	8	US-08-781-986A-518
C	28	39.6	2.1	3997	15	US-10-198-846-12639
C	29	39.6	2.1	6200	13	US-10-311-455-2187
C	30	39.4	2.1	484	11	US-09-918-995-30889
	31	39.4	2.1	1363	15	US-10-232-484-5
	32	39.4	2.1	2319	14	US-10-044-090-579
	33	39.4	2.1	3322	13	US-09-814-353-20277
C	34	39.4	2.1	11036	13	US-10-240-453-137
C	35	39.4	2.1	11036	15	US-10-239-676-117
	36	39	2.1	665	13	US-10-027-632-101927
	37	39	2.1	665	14	US-10-027-632-101927
C	38	38.8	2.1	487	10	US-09-960-352-7373
C	39	38.8	2.1	811	15	US-10-184-644-414
C	40	38.8	2.1	811	15	US-10-184-634-414
C	41	38.8	2.1	1187	13	US-10-027-632-101130
C	42	38.8	2.1	1187	14	US-10-027-632-101130
C	43	38.8	2.1	7928	13	US-10-311-455-67
C	44	38.8	2.1	3673778	13	US-10-312-841-1
C	45	38.6	2.1	516	13	US-09-814-353-17377

ALIGNMENTS

RESULT 1
US-09-745-763-35
Sequence 35, Application US/09745763
Patent No. US020065394A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallee, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Weisberg, David
Treacy, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 35:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1851 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 35:
 US-09-745-763-35

Query Match 100.0%; Score 1851; DB 9; Length 1851;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCGTAGCGCGGAGCTTACCTCTGGAGCCGCTCCGTCGCCGCTCAGAGCCGCTTA 60
DB 1 GCGTAGCGCGGAGCTTACCTCTGGAGCCGCTCCGTCGCCGCTCAGAGCCGCTTA 60
QY 61 TCAGATTATCTTAACAAGAAACCAACTGAAAAAATGAATTCCTTATCTTCGGAT 120
DB 61 TCAGATTATCTTAACAAGAAACCAACTGAAAAAATGAATTCCTTATCTTCGGAT 120
QY 121 TTTTCGGTGGTGTACCTTTTATCTCTGTGCTCTGGAAAGCTATATGCAAGATGCA 180
DB 121 TTTTCGGTGGTGTACCTTTTATCTCTGTGCTCTGGAAAGCTATATGCAAGATGCA 180
QY 181 TCTTAAGAGACCTTTTGAAGAAATTAAGAAATAGCCAGCTGTGGAGATGTTGCTA 240
DB 181 TCTTAAGAGACCTTTTGAAGAAATTAAGAAATAGCCAGCTGTGGAGATGTTGCTA 240
QY 241 AAGCAATCATCAACCTAGCTGTTTATGGTAAAGCCAGAACAGATCCTATGAGCATTTG 300
DB 241 AAGCAATCATCAACCTAGCTGTTTATGGTAAAGCCAGAACAGATCCTATGAGCATTTG 300
QY 301 CACTTCGTGTTGATACCTGTGAGCCAGACTGAGTGTCTCAAGAACTTAGAAAAAGCA 360
DB 301 CACTTCGTGTTGATACCTGTGAGCCAGACTGAGTGTCTCAAGAACTTAGAAAAAGCA 360
QY 361 TCCAATTATGTACCAAAACCTGACGCAAGATGGGCTGAGAAAGTTCACTGGAGCCAG 420
DB 361 TCCAATTATGTACCAAAACCTGACGCAAGATGGGCTGAGAAAGTTCACTGGAGCCAG 420
QY 421 TGAGAAATCCCACTGGAGAGGAGAGAAAGATCAGCTGTGATGCTGAGACCAAAATTC 480
DB 421 TGAGAAATCCCACTGGAGAGGAGAGAAAGATCAGCTGTGATGCTGAGACCAAAATTC 480
QY 481 ATAGATAGCCATCTGGGTCTTTGACAGACATTTGGACTCTCCAGAAAGCATTAAG 540
DB 481 ATAGATAGCCATCTGGGTCTTTGACAGACATTTGGACTCTCCAGAAAGCATTAAG 540
QY 541 CAGAAATCTGTGTGTGACCTTTTCGATGAACCTGACAGAAAGGGCTCAGAAAGCAAG 600
DB 541 CAGAAATCTGTGTGTGACCTTTTCGATGAACCTGACAGAAAGGGCTCAGAAAGCAAG 600
QY 601 GGAATATGTTGTTTATAACCAACCTTACATCACTCAAGACGGTGCAATACCGAA 660
DB 601 GGAATATGTTGTTTATAACCAACCTTACATCACTCAAGACGGTGCAATACCGAA 660
QY 661 CGCAGAGGGGCGGTGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTTGATCCGTGG 720
DB 661 CGCAGAGGGGCGGTGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTTGATCCGTGG 720
QY 721 CCTCTTCTTCATCTACAGTCTCTCAACAGTATTCAGGAATACCAAGATGGCGTCCA 780
DB 721 CCTCTTCTTCATCTACAGTCTCTCAACAGTATTCAGGAATACCAAGATGGCGTCCA 780
QY 781 AAATTCCAACGCGTATTAAGGTGGAAGATGCAAGAAATGATGTCAGAAATGGCTTCTC 840
DB 781 AAATTCCAACGCGTATTAAGGTGGAAGATGCAAGAAATGATGTCAGAAATGGCTTCTC 840
QY 841 ATGGGATCAAAATTTGTCAATTCAGCTTAAGATGGGGGCAAAAGACTACCCAGATCTGAT 900
DB 841 ATGGGATCAAAATTTGTCAATTCAGCTTAAGATGGGGGCAAAAGACTACCCAGATCTGAT 900

```

```

QY 901 CCTCAACACTGTAGCAGAGATCATCTGGAGCAAAATATCCAGAACAGTTGTACTGTGCA 960
DB 901 CCTCAACACTGTAGCAGAGATCATCTGGAGCAAAATATCCAGAACAGTTGTACTGTGCA 960
QY 961 GTGACATCTGACAGCTGGAGATGTTGGGAGGGTGCATGATGATGCGCGTGAAGCTT 1020
DB 961 GTGACATCTGACAGCTGGAGATGTTGGGAGGGTGCATGATGATGCGCGTGAAGCTT 1020
QY 1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTTGGGCTGCGTCCAAAGGACTC 1080
DB 1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTTGGGCTGCGTCCAAAGGACTC 1080
QY 1081 TGGGCTGTGTCTGTGACCTGACAGAAACAAAGGTGAGTTGGGCTTCAGATATATC 1140
DB 1081 TGGGCTGTGTCTGTGACCTGACAGAAACAAAGGTGAGTTGGGCTTCAGATATATC 1140
QY 1141 AGTTACACAAGGTAAATATTTCCAACTACAGTCTGTGATGAGTCTGACGCAAGAACT 1200
DB 1141 AGTTACACAAGGTAAATATTTCCAACTACAGTCTGTGATGAGTCTGACGCAAGAACT 1200
QY 1201 TCTTACCACTGGGCTGCAATTCATCTGGCACTGAAAGCCAGGGCCATCATGAGAGAG 1260
DB 1201 TCTTACCACTGGGCTGCAATTCATCTGGCACTGAAAGCCAGGGCCATCATGAGAGAG 1260
QY 1261 TTATGAGCTGTGACAGCCCTCCAAATATCACTCAGTCTGAGGCAATGAGAGAGAG 1320
DB 1261 TTATGAGCTGTGACAGCCCTCCAAATATCACTCAGTCTGAGGCAATGAGAGAGAG 1320
QY 1321 ACATCAACTTTTGGATCCAACTGAGTGTGCTGAGGCACTCTATGATGACTTATACA 1380
DB 1321 ACATCAACTTTTGGATCCAACTGAGTGTGCTGAGGCACTCTATGATGACTTATACA 1380
QY 1381 AGTATTTCTTCTTCATCATCTCCAGGAGACACATGACTGTATGATTCAAAGCAGA 1440
DB 1381 AGTATTTCTTCTTCATCATCTCCAGGAGACACATGACTGTATGATTCAAAGCAGA 1440
QY 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTTATGTTGTTGACAGACATGGAAGAA 1500
DB 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTTATGTTGTTGACAGACATGGAAGAA 1500
QY 1501 TGCTGCTAGGTCCTAGAAAACAGTAGAAGAAAGCTTTTCACTGCTTCTGGCCAGAAATC 1560
DB 1501 TGCTGCTAGGTCCTAGAAAACAGTAGAAGAAAGCTTTTCACTGCTTCTGGCCAGAAATC 1560
QY 1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCACTAACAATTTCAATTCATCTTCA 1620
DB 1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCACTAACAATTTCAATTCATCTTCA 1620
QY 1621 AAGCAAACTATTTTCAAGCTTCTGTATATCTTTCTGATGATCTTCCAAATCTCT 1680
DB 1621 AAGCAAACTATTTTCAAGCTTCTGTATATCTTTCTGATGATCTTCCAAATCTCT 1680
QY 1681 GATTCTGAAAAAAGAAATCATCTTCCCTCCCTCCCAACATAGAAATCAACATATGTA 1740
DB 1681 GATTCTGAAAAAAGAAATCATCTTCCCTCCCTCCCAACATAGAAATCAACATATGTA 1740
QY 1741 GGGATTACAGTGGGGGCAATTTCTTATATACCTCTTAAACCAATGTTTCCACTTTAAA 1800
DB 1741 GGGATTACAGTGGGGGCAATTTCTTATATACCTCTTAAACCAATGTTTCCACTTTAAA 1800
QY 1801 AGTAAACACTTAATTAATTTTGAAGATCTCTGAAAAAATGAAAAAATGAAAAA 1851
DB 1801 AGTAAACACTTAATTAATTTTGAAGATCTCTGAAAAAATGAAAAAATGAAAAA 1851

```

RESULT 2
 US-09-984-271-28
 ; Sequence 28, Application US/09984271
 ; Publication No. US20030040088A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 71 Human Secreted Proteins

FILE REFERENCE: P2030P1
 CURRENT APPLICATION NUMBER: US/09/984,271
 CURRENT FILING DATE: 2001-10-29
 PRIOR APPLICATION NUMBER: 09/482,273
 PRIOR FILING DATE: 2000-01-13
 PRIOR APPLICATION NUMBER: PCT/US99/15849
 PRIOR FILING DATE: 1999-07-14
 PRIOR APPLICATION NUMBER: 60/092,921
 PRIOR FILING DATE: 1998-07-15
 PRIOR APPLICATION NUMBER: 60/092,922
 PRIOR FILING DATE: 1998-07-15
 PRIOR APPLICATION NUMBER: 60/092,956
 PRIOR FILING DATE: 1998-07-15
 NUMBER OF SEQ ID NOS: 267
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 28
 LENGTH: 1863
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-984-271-28

Query Match 99.7%; Score 1846.2; DB 11; Length 1863;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1848; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 GGCTAGGCGGAGGCTTAGTCTCGGAGCCGCTCCGCGCGCGGCGAGAGCCGCCCTA 60
Db 1 GACTAGGCGGAGGCTTAGTCTCGGAGCCGCTCCGCGCGCGGCGAGAGCCGCCCTA 60
QY 61 TCAGATTATCTTAAACAAGAAACCACTGAAAGAAATTCCTTATCTTCCGAT 120
Db 61 TCAGATTATCTTAAACAAGAAACCACTGAAAGAAATTCCTTATCTTCCGAT 120
QY 121 TTTTCGGTGTGCTTCACTTTTATCCTGTGCTCTGSAAGATATATGCAAGATGCA 180
Db 121 TTTTCGGTGTGCTTCACTTTTATCCTGTGCTCTGSAAGATATATGCAAGATGCA 180
QY 181 TCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTGCTA 240
Db 181 TCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTGCTA 240
QY 241 AAGCAATCATCACTAGCTGTTTATGTTAAAGCCCAAGACGATCTATGAGCGATTGG 300
Db 241 AAGCAATCATCACTAGCTGTTTATGTTAAAGCCCAAGACGATCTATGAGCGATTGG 300
QY 301 CACTTCTGTGTGATCTGTGAGCCAGACTGAGTGTCTCAAGAACTTAAAGAAAGCCA 360
Db 301 CACTTCTGTGTGATCTGTGAGCCAGACTGAGTGTCTCAAGAACTTAAAGAAAGCCA 360
QY 361 TCCAAATTTATGATCAAAACCTGCGAGCAAGATGGGCTGGAAGAAATTCCTCGAGCCAG 420
Db 361 TCCAAATTTATGATCAAAACCTGCGAGCAAGATGGGCTGGAAGAAATTCCTCGAGCCAG 420
QY 421 TGAGATATCCCACTGAGAGAGGAGAAATCACTGTGATGTGAGCCAAAGATTC 480
Db 421 TGAGATATCCCACTGAGAGAGGAGAAATCACTGTGATGTGAGCCAAAGATTC 480
QY 481 ATAGATATGCACTCTGAGTCTTGTGAGCAAGATTTGGAATCTTCAAGAGGCAATTACAG 540
Db 481 ATAGATATGCACTCTGAGTCTTGTGAGCAAGATTTGGAATCTTCAAGAGGCAATTACAG 540
QY 541 CAGAAGTCTGTGTGATCTTCTTTCGATGAATCTGCAAGAGAGGCTTCAAGAGGCAAG 600
Db 541 CAGAAGTCTGTGTGATCTTCTTTCGATGAATCTGCAAGAGAGGCTTCAAGAGGCAAG 600
QY 601 GGAAGATTTGTTTATTAACCACTTACATCACTCAAGAGAGGCTGCAATACCGAA 660
Db 601 GGAAGATTTGTTTATTAACCACTTACATCACTCAAGAGAGGCTGCAATACCGAA 660
QY 661 CGCAGAGGAGGCTGAGAGCTGCAAGAGTGGGGCTTTGGCATCTCTCATTTGATCCGTGG 720
Db 661 CGCAGAGGAGGCTGAGAGCTGCAAGAGTGGGGCTTTGGCATCTCTCATTTGATCCGTGG 720

```

```

QY 721 CCTCCTTCTCATCTACAGTCTTCAACAGAGTATTCAGAAATACAGAGATGCGTCCCA 780
Db 721 CCTCCTTCTCATCTACAGTCTTCAACAGAGTATTCAGAAATACAGAGATGCGTCCCA 780
QY 781 AATTCCAAAGCCTGTATTAACGTTGGAAGATGCAAAATGATTCAGAAATGCTTCTC 840
Db 781 AATTCCAAAGCCTGTATTAACGTTGGAAGATGCAAAATGATTCAGAAATGCTTCTC 840
QY 841 ATGGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGACCTTACCAGATCTGAT 900
Db 841 ATGGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGACCTTACCAGATCTGAT 900
QY 901 CCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCA 960
Db 901 CCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCA 960
QY 961 GTGACATCTGACAGCTGGAGATGTTGGCAGAGGTCATGATGATGCGGTGAGCCT 1020
Db 961 GTGACATCTGACAGCTGGAGATGTTGGCAGAGGTCATGATGATGCGGTGAGCCT 1020
QY 1021 TTATATCATGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCGCCCAAGAGGACTC 1080
Db 1021 TTATATCATGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCGCCCAAGAGGACTC 1080
QY 1081 TGCGGCTGTGCTGTGACTGCAAGAAACAAGGTGAGTTGTGCTTCCAGTATATTC 1140
Db 1081 TGCGGCTGTGCTGTGACTGCAAGAAACAAGGTGAGTTGTGCTTCCAGTATATTC 1140
QY 1141 AGTTACAAGATTAATATTTCCACTACAGTCTGTGATGATGATGATGCGAGAACT 1200
Db 1141 AGTTACAAGATTAATATTTCCACTACAGTCTGTGATGATGATGATGCGAGAACT 1200
QY 1201 TCTTACCACTGCGCTGCAATTCACCTGCACTGAAAGGCGAGGCCATCATGAGAGAG 1260
Db 1201 TCTTACCACTGCGCTGCAATTCACCTGCACTGAAAGGCGAGGCCATCATGAGAGAG 1260
QY 1261 TTATGAGCTGTGCTGAGCCCTCAATATCACTCAAGTCTGAGCCATGAGAGAGGAG 1320
Db 1261 TTATGAGCTGTGCTGAGCCCTCAATATCACTCAAGTCTGAGCCATGAGAGAGGAG 1320
QY 1321 ACATCAACTTTTGGATCCAGACTGAGTCCGTGAGGCACTTATGATGATGATGATGAT 1380
Db 1321 ACATCAACTTTTGGATCCAGACTGAGTCCGTGAGGCACTTATGATGATGATGATGAT 1380
QY 1381 AGTATTTCTTCTTCACTCACTCCAGAGAGCACATGATGATGATGATGATGATGAT 1440
Db 1381 AGTATTTCTTCTTCACTCACTCCAGAGAGCACATGATGATGATGATGATGATGAT 1440
QY 1441 TGAATGTTGCTGCTGCTGTTTGGGCTGTGTTCTTATGTTGTGAGACATGAGAGAA 1500
Db 1441 TGAATGTTGCTGCTGCTGTTTGGGCTGTGTTCTTATGTTGTGAGACATGAGAGAA 1500
QY 1501 TGCTGCTAGGCTCTAGAAACAGTAAAGAAAGCTTTTATGATCTTCTGCGCAGAAATC 1560
Db 1501 TGCTGCTAGGCTCTAGAAACAGTAAAGAAAGCTTTTATGATCTTCTGCGCAGAAATC 1560
QY 1561 CTGGGTCTGCAACTTTGGAAGAACTCTTCACTCACTCACTCACTCACTCACTCACT 1620
Db 1561 CTGGGTCTGCAACTTTGGAAGAACTCTTCACTCACTCACTCACTCACTCACTCACT 1620
QY 1621 AAGCAACCTATTTATGATGCTTCTGATTAATCTTCTGATTAATCTTCTGATTAATCT 1680
Db 1621 AAGCAACCTATTTATGATGCTTCTGATTAATCTTCTGATTAATCTTCTGATTAATCT 1680
QY 1681 GATTTGAAAAAAGAAATCACTTCCCTCCCTCCCAACATGAAATCAATATGATGAT 1740
Db 1681 GATTTGAAAAAAGAAATCACTTCCCTCCCTCCCAACATGAAATCAATATGATGAT 1740
QY 1741 GGAATTAAGTGGGGGCAATTTCTTATATCACTCTTAAAGCAATGTTTCACTTTAA 1800
Db 1741 GGAATTAAGTGGGGGCAATTTCTTATATCACTCTTAAAGCAATGTTTCACTTTAA 1800
QY 1801 AGTAAACCTTAATTAATTTTGAAGATCTCTGAAAAAAGAAAAAAGAAAAA 1851

```

Db 1801 AGTAAACCTTATATATTTTGGAGAGATCTCTGAAAAA 1851

RESULT 3

US-09-917-800A-505

; Sequence 505, Application US/09917800A
; Patent No. US20020119462A1

GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 505
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AF097723
US-09-917-800A-505

Query Match 61.9%; Score 1145; DB 10; Length 1778;
Best Local Similarity 81.9%; Pred. No. 0;
Matches 1334; Conservative 0; Mismatches 290; Indels 5; Gaps 1;

QY 73 AACAGAAACCACTGAGAAAAAATGAATTCCTTATCTTCGATTTTGGTG 132
Db 95 AGCAAGAAAGAAAGAACTAGGACAAATGAGTTCCTTTCTCTGTTGCTGTTG 154
QY 133 TTCACCTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGSCATCTTAAGAGA 192
Db 155 TTCACCTTTTCTCCTTGGGCTCTGAAAAAGCTATATACAGAGTGTGTTTCTAGCGAA 214
QY 193 CTTTGAAAGAAATTAAGAAATAGCCAGCTGTGAGATTTGCTAAAGCAATCATCA 252
Db 215 CATTTCAAGAAATTAAGAAATAGCCAGCTATAGAGATGTGCTAAAGCAATTAATCA 274
QY 253 ACCTGCTGTTATGTAAGCCAGAAACAGATCCATAGAGGATTTGGACCTTCTGTTG 312
Db 275 ACCTGCTGTTATGTAAGAAATACAGAACCGATCGATAGAGCTTTGGACCTTCTGTTG 334
QY 313 ATACTGTGAGCCAGAGCTGAGTGGCTCCAGAAACCTAGAAAAAGCATCCAAATTATGT 372
Db 335 ATACTGTGAGCCAGAGCTGAGTGGCTCTTAAGAACTAGAAAGCTATCCAAATCATGT 394
QY 373 ACCAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGAGGCCAGTGAATATACCC 432

Db 395 ACCAAACCTGCAACAAGATGGCTGAAAAAGTCCACCTGAGAGCTGAGAAATACCTC 454
QY 433 ACTGGAGAGGGGAGAGAAATCAGCTGTGATGTGAGAGCCAAAGATTCATAGATAGCA 492
Db 455 ACTGGGAGGGGGAGAGAAATCTGATGATGTGTGCTCGAAATTCAGAAATTTGGCTA 514
QY 493 TCCTGGGTCTTGGAGAGAGATTTGGACCTCTCCAGAAAGGCAATTAAGAGAGTTCTG 552
Db 515 TTTTAGGCTTGGGCGGAGCATTTGGACTCTCTCGAAAGGTATACAGCAGAGACTCG 574
QY 553 TGTGACCTCTTTGATGTAATGAGAGAGGGCTCAGAGAGCAAGAGGAAATTTG 612
Db 575 TGTGAGCTCTTTGTTGTAATTCAGAAAGGGCATAGAGGCAAGAGGAAATTTG 634
QY 613 TTTATTAACCAACTTATACATCACTACTCAAGAGCGTGCAATACCAAGCGAGGGGCG 672
Db 635 TTTATTAACCACTTACCTGATCTATGAGAAACCTGTGCAATACCGGAGGCGGAGCTG 694
QY 673 TGAAGCTGCCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTTTCTCA 732
Db 695 TGAAGCTGCCAAGGTGGGGCGTGGCATCTCTCATTCGATCAGTAGCTTTTCTCA 754
QY 733 TCTACAGTCTTCAACAGATTTTCAAGATTCAGAGATGCGTCCCAAAATTTCCAAAG 792
Db 755 TCTACAGTCTTCAACAGATTCATCAAGATTCAGAGATGCGTCCCAAAATTTCCAAAG 814
QY 793 CCTGATTAACGAGTGAAGATGCAAGAAATGATGCAAGATGCTTCATCGATGAGATCAAA 852
Db 815 CCTGATTAACGATGAGATGCAAGAAATGATGCTTCGATGCTCTTCTGTGGGACAAA 874
QY 853 TTGTCAATTCAGCTAAAGATGGGGGCAAGACCTTACCCAGATTCGATTCCTTCAACCTG 912
Db 875 TTGTCAATTCAGCTAAAGATGGGGCAAGACCTTACCCAGATTCGATTCCTTCAACCTG 934
QY 913 TAGCAGATATCTGAGAGCAAAATTCAGAAACAGTTGTACTGTGATGAGATCTG 972
Db 935 TTGAGAGATATCTGAGAGCAAAATTCAGAAACAGTTGTACTGTGATGAGATCTG 994
QY 973 ACAGCTGGAGTGTGGGAGGGGCTCCAGATGATGAGTGGGAGCTTTATCATGAGG 1032
Db 995 ACAGCTGGAGTGTGGGAGGGGCTCCAGATGATGAGTGGGAGCTTTATCATGAGG 1054
QY 1033 AAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTTGGCGCTGCTG 1092
Db 1055 AAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTTGGCGCTGCTG 1114
QY 1093 TCTGAGCTGCAAGAAACAGAGTGTGCTGCTTCCATATTTATGATTAACAAG 1152
Db 1115 TCTGAGCTGCAAGAAACAGAGTGTGCTGCTTCCATATTTATGATTAACAAG 1174
QY 1153 TAAATATTTCCAACTACAGCTGTGATGAGTCTGAGCGCAGAACTTTTACCCACTG 1212
Db 1175 CAATATTTCCAACTACAGCTGTGATGAGTCTGAGCGCAGAACTTTTACCCACTG 1234
QY 1213 GGTGCAATTCAGCTGCAAGTAAAGGCGAGGCGCATCATGAGAGAGTTATGAGCTGC 1272
Db 1235 GGTGCAATTCAGCTGCAAGTAAAGGCGAGGCGCATCATGAGAGAGTTATGAGCTGC 1294
QY 1273 TGCAGCCCTCAATATCATCTAGGTCTGAGCCATGAGAGAGGACAGATCAACTTT 1332
Db 1295 TGCAGCCCTCAATATCATCTAGGTCTTATGATGAGAGAGGACAGATCAACTTT 1354
QY 1333 GGATTCAGAGTGAAGTCTGAGAGCTGACTGATGATGATGATGATGATGATGATGATG 1392
Db 1355 GGATTCAGAGTGAAGTCTGAGAGCTGACTGATGATGATGATGATGATGATGATGATG 1414
QY 1393 TCCATCACTCCCAAGAGACACATGACTGATGATGATGATGATGATGATGATGATG 1452
Db 1415 TCCATCACTCCCAAGAGACACATGACTGATGATGATGATGATGATGATGATGATG 1474
QY 1453 CTGCTGTTGGGCTGTTGTTTCTTATGTTGTTGAGATGAGAGAAATGCTGCTAGAGT 1512
Db 1475 CTGCTGTTGGGCTGTTGTTTCTTATGTTGTTGAGATGAGAGAAATGCTGCTAGAGT 1534

Oy	1513	CTTGAAGAACGTTTCAAGTCTGGCCAGGAATCTGGGCTGCA	15712
Db	1535	CCTAAGGAAACAGAGAGAGAACCTTGTCTTGCAGCTGGAGATCCCATTTGGGAT	15944
Oy	1573	CTTGGAAACCTCTTTCACATACAAATTCATCCAAATCTTCAAGACAACTCT	16322
Db	1595	TTTCACAGCAGCATCTTCAAGCAGCCTTGTATACATCAATCCCGTGGAC	16449
Oy	1633	ATTTCATGCTTTCTGTATATCTTTCTTGATACCTTCCAAATCTCTGATCTTAGAAA	16922
Db	1650	GTTTCTTATACCTTCTGTACCACTTTCCTTGATACGGCTTATACCTGTTCTTAGAATA	17099
Oy	1693	AGGATCAT	1701
Db	1710	AGTATCAT	1718

```

RESULT 4
US-09-984-271-95
; Sequence 95, Application US/09984271
; Publication No. US2003004008A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US/09/984,271
; PRIOR FILING DATE: 2000-01-13
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-984-271-95

```

Query Match	59.3%	Score 1098.4	DB 11	Length 1134
Best Local Similarity	99.7%	Pred. No. 0		
Matches 1121	Conservative	0	Mismatches 1	Indels 2
Gaps				2
QY	729	TCATCTTCAAGTCTTCAACAGGTATTGAGAAATACAGAGATGGCGTCCCAAAATTTCCA	788	
Db	1	TCATCTTCAAGTCTTCAACAGGTATTGAGAAATACAGAGATGGCGTCCCAAAATTTCCA	60	
OY	789	ACAGCTTGTATTTACGGTGGAGAGATGCAGAAATGATGTCAAGATGGCTTTTCATGGGATC	848	
Db	61	ACAGCTTGTATTTACGGTGGAGAGATGCAGAAATGATGTCAAGATGGCTTTTCATGGGATC	120	
OY	849	AAATTTGCATTCACACTTAAAGATGGGGGCAAAAGCCTTACCAGATACATGATTCCTTTCAAC	908	
Db	121	AAATTTGCATTCACACTTAAAGATGGGGGCAAAAGCCTTACCAGATACATGATTCCTTTCAAC	180	
OY	909	ACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAAACAGTTGTACTGGTCAGTGGACAT	968	
Db	181	ACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAAACAGTTGTACTGGTCAGTGGACAT	240	
OY	969	CTGACAGCTGGAGTGTGGCAGGGTGCATGATGATGCGGTGAGACCTTTATATCA	1028	
Db	241	CTGACAGCTGGAGTGTGGCAGGGTGCATGATGATGCGGTGAGACCTTTATATCA	300	
OY	1029	TGGGAAGCACTTCATTTATTTAAAGATCTTTGGGCTGGCCCAAAAGACCTTTGGGGCTG	1088	
Db	301	TGGGAAGCACTTCATTTATTTAAAGATCTTTGGGCTGGCCCAAAAGACCTTTGGGGCTG	360	

OY	1089	GTGCTCGGACCTGAGAAAGAAACAAGGGTGGAGTTGGTGCTTCCAGATATTAACGATTAC	1148
Db	361	GTGCTCTGGACATGAGAAAGAAACAAGGGAGATTGGTGCTTCCAGATATTAACGATTAC	420
OY	1149	AAGSTAAATATTTTCCAACTACAGTCTGTGTGATGAGTCTGACGAGAACTTTCTAAC	1208
Db	421	AAGSTAAATATTTTCCAACTACAGTCTGTGTGATGAGTCTGACGAGAACTTTCTAAC	480
OY	1209	ACTGGGCTGCATTTCACTGGCAGTGAAGAAAGGCCATCATGAGAGAGTTATGAGC	1268
Db	481	ACTGGGCTGCATTTCACTGGCAGTGAAGAAAGGCCATCATGAGAGAGTTATGAGC	539
OY	1269	CTGCTGCAGCCCTCCATPATCACTCAGGTCCGTGAGCATGGAGAAAGGACAGACATCAAC	1327
Db	540	CTGCTGCAGCCCTCCATPATCACTCAGGTCCGTGAGCATGGAGAAAGGACAGACATCAAC	599
OY	1329	TTTTGATTCACAGCTGAGTGGCTTGAGCCAGTCTACTTGATGACTTATACAGATATTC	1388
Db	600	TTTTGATTCACAGCTGAGTGGCTTGAGCCAGTCTACTTGATGACTTATACAGATATTC	659
OY	1389	TTCTTTCCATCACTCTCCACGAGACACCATGACTGTCAATGATTCGAAGCAGATGAATGT	1448
Db	660	TTCTTTCCATCACTCTCCACGAGACACCATGACTGTCAATGATTCGAAGCAGATGAATGT	719
OY	1449	GGTGTGTGTGTTTGGGCTGTGTTTCTTAATGTGTGAGACATGGAGAAATGTGCT	1508
Db	720	GGTGTGTGTGTTTGGGCTGTGTTTCTTAATGTGTGAGACATGGAGAAATGTGCT	779
OY	1509	AGGTCTAGAAACAAGTAAGAAAGAAAGCTTTTCAATGCTTGGCCAGGAACTCTGGATCT	1568
Db	780	AGGTCTAGAAACAAGTAAGAAAGAAAGCTTTTCAATGCTTGGCCAGGAACTCTGGATCT	839
OY	1569	GCAACTTTGGAAAACTCCTCTTTCACATTAACAATTTCAATCCAAATTCATTTCAAGACAA	1627
Db	840	GCAACTTTGGAAAACTCCTCTTTCACATTAACAATTTCAATCCAAATTCATTTCAAGACAA	899
OY	1629	CTCATATTTCAAGCTTTCTGTATTAATCTTTCTGTATTAATTTTCAAAATTCCTCTG-ATTCTA	1687
Db	900	CTCATATTTCAAGCTTTCTGTATTAATCTTTCTGTATTAATTTTCAAAATTCCTCTG-ATTCTA	959
OY	1688	GAAGAAAGAAATCATTTCTCCCTCCCTCCGACACATAGAAATCAATATGTGTAGGATTA	1747
Db	960	GAAGAAAGAAATCATTTCTCCCTCCCTCCGACACATAGAAATCAATATGTGTAGGATTA	1011
OY	1748	CAGTGGGGGCACTTTCTTTATATACACTTTTAAAAACATTTGTTTCCACTTTAAAGTAAC	1807
Db	1020	CAGTGGGGGCACTTTCTTTATATACACTTTTAAAAACATTTGTTTCCACTTTAAAGTAAC	1073
OY	1808	ACTTAATATAATTTTGGAGAGATCTGTGAAAAAATTTTTTTTTTTTAAAAA 1851	
Db	1080	ACTTAATATAATTTTGGAGAGATCTGTGAAAAAATTTTTTTTTTTTAAAAA 1123	

```

RESULT 5
US-09-833-381-1929
; Sequence 1929, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1929
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```


CURRANT FILING DATE: 2001-07-20
 PRIOR APPLICATION NUMBER: US 60/287,722
 PRIOR FILING DATE: 2001-05-02
 PRIOR APPLICATION NUMBER: US 60/221,607
 PRIOR FILING DATE: 2000-07-28
 NUMBER OF SEQ ID NOS: 32337
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 8960
 LENGTH: 60

1 APPLICANT: Yamamoto, Jun-ichi
2 APPLICANT: Ishii, Shizuko
3 APPLICANT: Sugiyama, Tomoyasu
4 APPLICANT: Wakamatsu, Ai
5 APPLICANT: Nagai, Keiichi
6 APPLICANT: Otsuki, Tetsuji
7 APPLICANT: Funahashi, Shin-Ichi
8 APPLICANT: Senoo, Chiaki
9 APPLICANT: Nezu, Jun-ichi

TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-098001
; CURRENT APPLICATION NUMBER: US/10/059,585
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)...(1450)
US-10-059-585-5

Query Match 2.3%; Score 42; DB 15; Length 2200;
Best Local Similarity 45.5%; Pred. No. 0.49;
Matches 150; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

Qy 933 AAATATCCAGAACAGGTTGTACTGTGTCAGTCAATCGACATCGACAGCTGGAGTTGGGAG 992
Db 158 AAGGATGAGCAGGAGATGTGTCAGAGGAGAGAGCTGCTGAGTTTCATGGGAGGAG 217
Qy 993 GGTGCATGATGATGAGCGGTGAGCCCTTTATCATGGAAGCACTCTCACTTATTAA 1052
Db 218 GCAGCCCTGACCGACGCGAGGTGGCCGGAGAGGGGTGGCTGGCCCTCTTCAAGG 277
Qy 1053 GATCTGGGCTGCGTCCAAAGAGACTCTGCGGCTGGTGTCTGTGAGCTGCAGAAACAA 1112
Db 278 GAGAGAGGGCGACCTCACTGTCAGAGAGTGTGCTGCTCTTCAAGTGAAGAGAGAA 337
Qy 1113 GGTGAGTGTGCTGCTTCAATTTATCATGTTACAGAAAGTAAATTTCCACTACAGT 1172
Db 338 GAGAGAGAGAGTGTGCTGAGTCTTGTACCAACCAAGGCTCTGCGCCCGACTCAGC 397
Qy 1173 CTGGTATGAGTGTGACGACGAACTTTTACCCACTGGGCTGCAATTCATCTGGCAGT 1232
Db 398 ATCCCTGCTCTTATATCAGACGTCGGACAAACACAGCCCTGAGATGCTGACCGG 457
Qy 1233 GAAAAGGCCAGGGCCATCATGAGAGGTT 1262
Db 458 GAGAACACGTTTGTGTAAAGTGAAGT 487

RESULT 15
US-10-311-455-1458/c
; Sequence 1458, Application US/10311455
; Publication No. US20030143606a1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1458
; LENGTH: 13202
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1458

Query Match 2.3%; Score 41.8; DB 13; Length 13202;
Best Local Similarity 49.3%; Pred. No. 2;
Matches 109; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 1631 CTATTCATGCTTCTGTTATATCTTTGTGATCTTCCAAATCTCTGATTCTAGAA 1690
Db 3058 CCAATTCCTCCATCTTAAATTTCTTAAATACCACTTAAATTAACCAATTCATTA 2999
Qy 1691 AAAGATATCTTCCCTCCCTCCACACATAGATCAATATATGATGAGTACAG 1750
Db 2998 AATCTAAATATCTTCTTCTTACCTTCAATCTTAAATATCTGAAATAAATTAAT 2939
Qy 1751 TGGGGCATTTCTTATATCACTCTTAAACAAATGTTTCACTTAAAGTAAACACT 1810
Db 2938 TATTAATATATCTTAAACAACTTACATCACTTACCTTCAACATCTTAAACCC 2879
Qy 1811 TAAATAAATTTTGGAGATCTTGAAAAAATTTTAAAAAATTTTAAAAAATTTT 1851
Db 2878 TAAATATTAATCTTAAAAAATCTTATATATATTAATTAATTAATTAATTAAT 2838

Search completed: December 22, 2003, 23:44:36
Job time : 589.358 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:18:08 ; Search time 5111.6 Seconds

(without alignments)
11332.658 Million cell updates/sec

Title: US-09-745-763-35_COPY_99_1514

Perfect score: 1416

Sequence: 1 ATGAATTCCTTATCTTCTGC.....AAGAAATGCTGCTAGCTCC 1416

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl1.*

1: gb Da.*

2: gb Hcg.*

3: gb In.*

4: gb Om.*

5: gb Ov.*

6: gb Pat.*

7: gb Ph.*

8: gb Pl.*

9: gb Pr.*

10: gb Ro.*

11: gb Sts.*

12: gb Sy.*

13: gb Un.*

14: gb Vi.*

15: em Da.*

16: em Fun.*

17: em Hum.*

18: em In.*

19: em Mu.*

20: em Om.*

21: em Or.*

22: em Ov.*

23: em Pat.*

24: em Ph.*

25: em Pl.*

26: em Ro.*

27: em Sts.*

28: em Un.*

29: em Vi.*

30: em Hcg Hum.*

31: em Hcg Inv.*

32: em Hcg Other.*

33: em Hcg Mus.*

34: em Hcg Pln.*

35: em Hcg Rod.*

36: em Hcg Mam.*

37: em Hcg Vrt.*

38: em Gy.*

39: em Htgo Hum.*

40: em Htgo Mus.*

41: em Htgo Other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1416	100.0	1851	6	BD106411 Secreted
2	1412.8	99.8	1778	6	BD105816 Secretory
3	1412.8	99.8	1860	6	AX006440 Sequence
4	1412.8	99.8	1928	9	BC020689 Homo sapi
5	1409.6	99.5	1472	9	AF107834 Homo sapi
6	1408	99.4	1860	6	BD127520 Primer fo
7	1408	99.4	1860	6	AK075132 Homo sapi
8	1399.2	98.8	1794	9	AF119386 Homo sapi
9	1116.8	78.9	1796	10	BC037067 Mus muscu
10	1115.2	78.8	1596	10	AF107835 Mus muscu
11	1107.2	78.2	1726	10	AF131077 Rattus no
12	1107.2	78.2	1778	6	AX400829 Sequence
13	1107.2	78.2	1778	10	AF097723 Rattus no
14	1030.8	72.8	1716	10	AF009513 Mus muscu
15	520.6	36.8	609	6	BD059610 Secreted
16	470.2	33.2	642	6	BD125210 Primer fo
17	470.2	33.2	642	6	BD126476 Primer fo
18	436.4	30.8	156702	2	AC084223 Homo sapi
19	436.4	30.8	157927	9	AP006278 Homo sapi
20	436.4	30.8	171301	9	AC010859 Homo sapi
21	421.4	29.8	423	9	AF107833 Homo sapi
22	331	23.4	417	10	AF107832 Mus muscu
23	325.8	23.0	167883	2	AC121026 Rattus no
24	325.8	23.0	218874	2	AC110103 Rattus no
25	323.2	22.8	214287	2	AC128951 Mus muscu
26	284.8	20.1	314	6	BD058492 Secreted
27	276.6	19.5	462	6	BD105811 Secretory
28	276.6	19.5	462	6	E62849 Secretary P
29	212.6	15.0	317	6	BD076934 5' EST of
30	212.2	15.0	144703	9	AP003112 Homo sapi
31	205	14.5	152176	9	AP003117 Homo sapi
32	205	14.5	171035	9	AP003111 Homo sapi
33	174.4	12.3	214287	2	AC128951 Mus muscu
34	172.8	12.2	203469	2	AC123654 Mus muscu
35	171.2	12.1	218874	2	AC110103 Rattus no
36	163.4	11.5	166050	2	AC013817 Homo sapi
37	163.4	11.5	177008	2	AC011175 Homo sapi
38	163.4	11.5	240178	2	AC094173 Rattus no
39	163.4	11.5	288136	2	AC120489 Rattus no
40	154.4	10.9	238034	2	AC123681 Mus muscu
41	116	8.2	10792	1	AE015790 Shewanell
42	115	8.1	110000	2	AL390202_01 Continuation (2 of
43	114.2	8.1	492	6	BD125961 Primer fo
44	99.2	7.0	10487	1	AE005922 Caulobact
45	96.4	6.8	1729	3	AF077194 Acanthoch

ALIGNMENTS

RESULT 1
BD106411
LOCUS BD106411 1851 bp DNA linear PAT 18-SEP-2002
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD106411
VERSION BD106411.1 GI:23201229
KEYWORDS JP 2002503955-A/2.
SOURCE Chlamydia sp.
ORGANISM Chlamydia sp.
REFERENCE 1 (bases 1 to 1851)
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Werberg,D.,
Treacy,M., Spaulding,V. and Agostino,M.J.
TITLE Secreted proteins and polynucleotides encoding them
JOURNALS Patent: JP 2002503955-A 2 05-FEB-2002;

Pred. No. is the number of results predicted by chance to have a

COMMENT

GENETICS INSTITUTE INC
 PN JP 2002503955-A/2
 PD 05-FEB-2002
 PF 20-MAR-1998 JP 1998545874
 PR 21-MAR-1997 US 08/822167, 19-MAR-1998 US 09/044466 PI
 KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLE, LISA A RACIE, PI
 DAVID MERBERG,
 PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
 C12N15/12, C07K14/47, A61K38/17
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key location/Qualifiers.

FEATURES
 source location/Qualifiers
 1. 1851
 /organism="Chlamydia sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:35827"

BASE COUNT 531 a 413 c 438 g 469 t

ORIGIN

Query Match 100.0%; Score 1416; DB 6; Length 1851;
 Query Similarity 100.0%; Pred. No. 0;
 Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAATTCCTATCTTGCATTTTCCGTGGTTCACCTTTATCCCTGCTGGG 60
 99 ATGAATTCCTATCTTGCATTTTCCGTGGTTCACCTTTATCCCTGCTGGG 158
 61 AAAGCTATATGCAAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATA 120
 159 AAAGCTATATGCAAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATA 218
 121 GCCAGCTGTGAGAGTGTGCTTAAGACATCATCAACCTGCTTTATGTAAGAGCCAG 180
 219 GCCAGCTGTGAGAGTGTGCTTAAGACATCATCAACCTGCTTTATGTAAGAGCCAG 278
 181 AACAGATCTTATGAGCGATTTGGCACTTGTGTTGATCTGTGAGACCCAGACTGAGTGC 240
 279 AACAGATCTTATGAGCGATTTGGCACTTGTGTTGATCTGTGAGACCCAGACTGAGTGC 338
 241 TCCAAGAACTTAAGAAAAAGCCATCCAAATTAATGTACAAAACCTGACCAAGATGGGCTG 300
 339 TCCAAGAACTTAAGAAAAAGCCATCCAAATTAATGTACAAAACCTGACCAAGATGGGCTG 398
 301 GAGAAAGTTCACTGAGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAAGATCAACT 360
 399 GAGAAAGTTCACTGAGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAAGATCAACT 458
 361 GTGATGTGAGAGCCAAATTCATTAAGATAGCCATCTGGGTCTTGGAGAGCACTTGGG 420
 459 GTGATGTGAGAGCCAAATTCATTAAGATAGCCATCTGGGTCTTGGAGAGCACTTGGG 518
 421 ACTCTTCAGAAAGCATTACAGACAGATTTCTGTGTGATCCTTTTGCATGAATCTGAG 480
 519 ACTCTTCAGAAAGCATTACAGACAGATTTCTGTGTGATCCTTTTGCATGAATCTGAG 578
 481 AGAAGGGCTTCAGAAAGCAGAGGAGAGATGTTGTTTAAACCACTTATCACTACATAC 540
 579 AGAAGGGCTTCAGAAAGCAGAGGAGAGATGTTGTTTAAACCACTTATCACTACATAC 638
 541 TCAAGAGCGTGCATATCCGAACGAGGGGGGCTGGAAGCTGCCAAGTGGGGCTTTG 600
 639 TCAAGAGCGTGCATATCCGAACGAGGGGGGCTGGAAGCTGCCAAGTGGGGCTTTG 698
 601 GCATCTCTCATTCGATTCGCTGGCTCTTCTTCATCTACAGTCTCTACACAGATATTCAG 660
 699 GCATCTCTCATTCGATTCGCTGGCTCTTCTTCATCTACAGTCTCTACACAGATATTCAG 758
 661 GAATATCCAGAGATGGGTGCCCAAAATTTCCACAGCTGTATTAAGGTGAAGATGCAAGA 720
 759 GAATATCCAGAGATGGGTGCCCAAAATTTCCACAGCTGTATTAAGGTGAAGATGCAAGA 818
 721 ATGATGTCAAGAAATGGCTTCTCATGTGATCAAAATTTGTATTCAGCTAAAGATGGGGCA 780

819 ATGATGTCAAGAAATGGCTTCTCATGTGATCAAAATTTCTATTCAGTGAAGATGGGGCA 878
 781 AAGACCTTACCAAGATATGATTTCTTCAACACTGTAGACAGATCATCTGGAGCAATAT 840
 879 AAGACCTTACCAAGATATGATTTCTTCAACACTGTAGACAGATCATCTGGAGCAATAT 938
 841 CCAAGACAGTTGATCTGATGATGACATCTGGAACAGCTGGAGATTTGGGAGAGTGC 900
 939 CCAAGACAGTTGATCTGATGATGACATCTGGAACAGCTGGAGATTTGGGAGAGTGC 998
 901 ATGATGATGAGCGGTGAGGCTTTTATATCATGGAAGCACTCTCATTTATTAAGATCTT 960
 999 ATGATGATGAGCGGTGAGGCTTTTATATCATGGAAGCACTCTCATTTATTAAGATCTT 1058
 961 GGGCTGTGTCCAAGAGACTCTGCGGCTGTGTCTGTGACTGACAGAGAAAGAGTGA 1020
 1059 GGGCTGTGTCCAAGAGACTCTGCGGCTGTGTCTGTGACTGACAGAGAAAGAGTGA 1118
 1021 GTTGTGCTCTCCAGATATATCATGATTTACAGAGTAATATTTCCAACTACAGTCTGTG 1080
 1119 GTTGTGCTCTCCAGATATATCATGATTTACAGAGTAATATTTCCAACTACAGTCTGTG 1178
 1081 ATGAGATCTGACGAGAACCTTTTACCACTGGGCTGCAATTCATGSCAGTGAAG 1140
 1179 ATGAGATCTGACGAGAACCTTTTACCACTGGGCTGCAATTCATGSCAGTGAAG 1238
 1141 GCCAGGGCCATCATGAGAGAGTTATAGCTGTGAGAGCCCTTCATATCATCACTAGTC 1200
 1239 GCCAGGGCCATCATGAGAGAGTTATAGCTGTGAGAGCCCTTCATATCATCACTAGTC 1298
 1201 CTGAGCCATGGAAGAGGAGACAGACATCACTTTTGAATCCAAAGCTGAGTGCCTGAGCC 1260
 1299 CTGAGCCATGGAAGAGGAGACAGACATCACTTTTGAATCCAAAGCTGAGTGCCTGAGCC 1358
 1261 AGTCTACTGATGATCTTATCAAGATATTTCTTCCATGACCTGCCAGAGACCATG 1320
 1359 AGTCTACTGATGATCTTATCAAGATATTTCTTCCATGACCTGCCAGAGACCATG 1418
 1321 ACTGTCAATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1380
 1419 ACTGTCAATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1478
 1381 GTTGTGACAGACATGGAAGAAATGCTGCTAGTGC 1416
 1479 GTTGTGACAGACATGGAAGAAATGCTGCTAGTGC 1514

RESULT 2
 BD105816
 LOCUS
 DEFINITION
 SECRETORY PROTEIN.
 ACCESSION
 BD105816.1 GI:23200634
 VERSION
 JP 2002502234-A/7.
 KEYWORDS
 SYNTHETIC CONSTRUCT
 SOURCE
 ORGANISM
 SYNTHETIC CONSTRUCT
 artificial sequences.
 1 (bases 1 to 1778)
 JACOBS, K., MCCOY, J. M., RACIE, L. A., LAVALLE, B. R., MERBERG, D. and
 SPAULDING, V.
 Secretory protein.
 Patent: JP 2002502234-A 7 22-JAN-2002;
 TITLE
 JOURNAL
 COMMENT

GENETICS INSTITUTE INC
 PN JP 2002502234-A/7
 PD 22-JAN-2002
 PF 16-APR-1997 JP 1997537384
 PR 18-APR-1996 US 08/634325, 13-JAN-1997 US 08/783520 PI
 KENNETH JACOBS, JOHN M MCCOY, LISA A RACIE, EDWARD R LAVALLE, PI
 DAVID MERBERG,
 PI VIKKI SPAULDING
 PC C12N15/12, C07K14/47, A61K38/17
 CC Strandedness: Double;

CC Topology: Linear; Location/Qualifiers.
 FH Key Location/Qualifiers
 1. 1778
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

BASE COUNT 514 a 386 c 422 g 456 t
 ORIGIN

Query Match 99.8%; Score 1412.8; DB 6; Length 1778;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 ATGAAATTCCTTATCTCGCATTTTGGTGTGACCTTTATCCCTGTGCTGGG 60
Db 17 ATGAAATTCCTTATCTCGCATTTTGGTGTGACCTTTATCCCTGTGCTGGG 76
Qy 61 AAAGCTATATGCAAGAAATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAAATA 120
Db 77 AAAGCTATATGCAAGAAATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAAATA 136
Qy 121 GCCAGCTGTGAGATGTTGCTAAAGCAATCAACTAGCTTTTATGTAAGCCAG 180
Db 137 GCCAGCTGTGAGATGTTGCTAAAGCAATCAACTAGCTTTTATGTAAGCCAG 196
Qy 181 AACGATCTATGAGCGATGGGCACTTGTGTATATCTGTGTAACCGACGTAAGTGC 240
Db 197 AACGATCTATGAGCGATGGGCACTTGTGTATATCTGTGTAACCGACGTAAGTGC 256
Qy 241 TCCAAAGAACTAGAAAAAGCCATCCAAATTAATGTAACAAAACCTGACAGATGGGCTG 300
Db 257 TCCAAAGAACTAGAAAAAGCCATCCAAATTAATGTAACAAAACCTGACAGATGGGCTG 316
Qy 301 GAGAAAGTTCACTGAGAGCCAGTAAGTAATCCCACTGGAGAGAGAGAAATCAAGCT 360
Db 317 GAGAAAGTTCACTGAGAGCCAGTAAGTAATCCCACTGGAGAGAGAGAAATCAAGCT 376
Qy 361 GTGATGCTGAGAGCCAAATTCATAGATAGCCATCTGGGTCTTTGGCAGAGATGGG 420
Db 377 GTGATGCTGAGAGCCAAATTCATAGATAGCCATCTGGGTCTTTGGCAGAGATGGG 436
Qy 421 ACTCTCTCAGAAAGCAATTACAGAGAAATTTCTGGTGTGACCTCTTTCAATGAATCGAG 480
Db 437 ACTCTCTCAGAAAGCAATTACAGAGAAATTTCTGGTGTGACCTCTTTCAATGAATCGAG 496
Qy 481 AGAAGGCTCTCAGAAAGCAAGAGAGATGTTGTTATTAACAACCTTACATCACTAC 540
Db 497 AGAAGGCTCTCAGAAAGCAAGAGAGATGTTGTTATTAACAACCTTACATCACTAC 556
Qy 541 TCAAGAGCGGTGCAATACCGAAGCAGAGGGGGGTGGAAGCTGCCAAGGTGGGGCTTTG 600
Db 557 TCAAGAGCGGTGCAATACCGAAGCAGAGGGGGGTGGAAGCTGCCAAGGTGGGGCTTTG 616
Qy 601 GCATCTCTCAATCGATCGGTGGCTCTCTTCATCTCAATCAAGTCTCAACAGATATTAG 660
Db 617 GCATCTCTCAATCGATCGGTGGCTCTCTTCATCTCAATCAAGTCTCAACAGATATTAG 676
Qy 661 GAATATCAGAGATGGCGTCCCAAAATTCACAAGCTGTATTAAGTGTGAGATGAGAA 720
Db 677 GAATATCAGAGATGGCGTCCCAAAATTCACAAGCTGTATTAAGTGTGAGATGAGAA 736
Qy 721 ATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGTCATTCAGCTTAAGATGGGGCA 780
Db 737 ATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGTCATTCAGCTTAAGATGGGGCA 796
Qy 781 AAAGCTACCAAGATGATGATCTTCTTCAACCTGTAGAGAGATCACTGGGAGCAATAT 840
Db 797 AAAGCTACCAAGATGATGATCTTCTTCAACCTGTAGAGAGATCACTGGGAGCAATAT 856
Qy 841 CCAGAACAGGTTGATGTGTCAGTGAATCTGGAACAGCTGGAGTGTGGCAGAGGTGCC 900
Db 857 CCAGAACAGGTTGATGTGTCAGTGAATCTGGAACAGCTGGAGTGTGGCAGAGGTGCC 916
  
```

Qy 901 ATGATGATGCGGCTGAGGCTTTATATCATGGAGACACTCTGACTTATTAAGATCTT 960
 Db 917 ATGATGATGCGGCTGAGGCTTTATATCATGGAGACACTCTGACTTATTAAGATCTT 976
 Qy 961 GGGCTGCGTCCAAAGAGACTCTGCGGCTGGTGTCTGTAAGTGAAGAAAGAGTGA 1020
 Db 977 GGGCTGCGTCCAAAGAGACTCTGCGGCTGGTGTCTGTAAGTGAAGAAAGAGTGA 1036
 Qy 1021 GTTGTGCTTCCAGTATTAATCACTTACACAGGTAATATTTTCAACTACAGTGTG 1080
 Db 1037 GTTGTGCTTCCAGTATTAATCACTTACACAGGTAATATTTTCAACTACAGTGTG 1096
 Qy 1081 ATGAGCTGACGAGGAGACCTTCTTACCACTGGGCTGGAATTCATCGAGAGAAAG 1140
 Db 1097 ATGAGCTGACGAGGAGACCTTCTTACCACTGGGCTGGAATTCATCGAGAGAAAG 1156
 Qy 1141 GCCAGGCAATCATGAGAGAGGTTATAGGCTGTGACGCCCTCAATATCACTCAGGTC 1200
 Db 1157 GCCAGGCAATCATGAGAGAGGTTATAGGCTGTGACGCCCTCAATATCACTCAGGTC 1216
 Qy 1201 CTGAGCCATGAGAGAGGACACATCACTTTTGAATCCAAGCTGAGTGCCTGAGCC 1260
 Db 1217 CTGAGCCATGAGAGAGGACACATCACTTTTGAATCCAAGCTGAGTGCCTGAGCC 1276
 Qy 1261 AGTCTACTGATGATCACTTATCAAGTATTTCTTCCATCATCTCCACGAGACCATG 1320
 Db 1277 AGTCTACTGATGATCACTTATCAAGTATTTCTTCCATCATCTCCACGAGACCATG 1336
 Qy 1321 ACTGTCAATGATCCAAAGAGATGAATGTTGCTGCTGCTGTTGGGCTGTTCTTAT 1380
 Db 1337 ACTGTCAATGATCCAAAGAGATGAATGTTGCTGCTGCTGTTGGGCTGTTCTTAT 1396
 Qy 1381 GTTGTGCAACATGAGAAAGATGCTGCTGAGTGC 1416
 Db 1397 GTTGTGCAACATGAGAAAGATGCTGCTGAGTGC 1432

RESULT 3
 AX006440 1860 bp DNA linear PAT 06-SEP-2000
 LOCUS
 DEFINITION Sequence 5 from Patent WO0004157.
 AX006440
 ACCESSION AX006440
 VERSION AX006440.1 GI:9994575
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Neeff, J.M., Peeters, D.C. and Pangalos, M.
 Cloning and characterisation of novel mammalian peptidases
 Patent: WO 0004157-A 5 27-JAN-2000;
 JANSSEN PHARMACEUTICA NV (BE); NEEFS JEAN MARC EDMOND FERNAND (BE);
 PETERS DANIELLE CELINE GEORGE (BE); PANGALOS MERLIS (GB)

FEATURES
 source
 1. 1860
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 518 a 420 c 451 g 471 t
 ORIGIN

Query Match 99.8%; Score 1412.8; DB 6; Length 1860;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 ATGAAATTCCTTATCTTGGCATTTTGGTGTGACCTTTATCCCTGTGCTGGG 60
Db 125 ATGAAATTCCTTATCTTGGCATTTTGGTGTGACCTTTATCCCTGTGCTGGG 184
Qy 61 AAAGCTATATGCAAGAAATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAAATA 120
  
```


Db 185 AAGCTATATGCAAGATGCAATCTCTTAAGAGGACTTTTGAAGAAATTAAGAAATA 244
 Qy 121 GCGAGCTGTGAGATGTTGCTAAAGCATTCATCACTAGCTTTTATGTTAAAGCCAG 180
 Db 245 GCGAGCTGTGAGATGTTGCTAAAGCATTCATCACTAGCTTTTATGTTAAAGCCAG 304
 Qy 181 AACAGATCTATGAGCGATTTGGCACTTCTGTTGATACGTGTGACCCAGACTGAGTGGC 240
 Db 305 AACAGATCTATGAGCGATTTGGCACTTCTGTTGATACGTGTGACCCAGACTGAGTGGC 364
 Qy 241 TCCAAGAACTTAAGAAAAAGCCATTCATTAATTAATTAACAAAACCTGACAGCAAGATGGCTG 300
 Db 365 TCCAAGAACTTAAGAAAAAGCCATTCATTAATTAATTAACAAAACCTGACAGCAAGATGGCTG 424
 Qy 301 GAGAAAGTTCACTGAGACCGAGTGAATACCCCACTGGAGAGAGGGAGAAAGATCACT 360
 Db 425 GAGAAAGTTCACTGAGACCGAGTGAATACCCCACTGGAGAGAGGGAGAAAGATCACT 484
 Qy 361 GTGATGCTGAGACCAAGAAATTCATTAAGATAGCCATCTGGGCTTTGGCAGCAGATTGGG 420
 Db 485 GTGATGCTGAGACCAAGAAATTCATTAAGATAGCCATCTGGGCTTTGGCAGCAGATTGGG 544
 Qy 421 ACTCTCCAGAAAGCATTAACAGCAGAGTTCTGTGTGACCTCTTTGATGAATGCGAG 480
 Db 545 ACTCTCCAGAAAGCATTAACAGCAGAGTTCTGTGTGACCTCTTTGATGAATGCGAG 604
 Qy 481 AGAAGGGCTCTGAGAAAGCAAGAGAGAGTGTGTTTAACCAACCTTATCACTCACTAC 540
 Db 605 AGAAGGGCTCTGAGAAAGCAAGAGAGAGTGTGTTTAACCAACCTTATCACTCACTAC 664
 Qy 541 TCAAGAGCTGTCATATACCAAGCAGAGGGGGCGGTGAGAGCTGCCAAGTGGGGCTTTG 600
 Db 665 TCAAGAGCTGTCATATACCAAGCAGAGGGGGCGGTGAGAGCTGCCAAGTGGGGCTTTG 724
 Qy 601 GCATCTCTCATTCGATCCGTGGCTCTTCTTCATCTCAAGTCTCACAAGATTAATGAG 660
 Db 725 GCATCTCTCATTCGATCCGTGGCTCTTCTTCATCTCAAGTCTCACAAGATTAATGAG 784
 Qy 661 GAATACCAAGATGAGGGTCCCAAAATTCACAAGCTGATTAACGTTGGAAGATGCAAA 720
 Db 785 GAATACCAAGATGAGGGTCCCAAAATTCACAAGCTGATTAACGTTGGAAGATGCAAA 844
 Qy 721 ATGATGTCAGAAAGTGGCTTCTTCATGAGATCAAAATTTGATTCAGCTAAAGTGGGGCA 780
 Db 845 ATGATGTCAGAAAGTGGCTTCTTCATGAGATCAAAATTTGATTCAGCTAAAGTGGGGCA 904
 Qy 781 AAGACTTACCAAGATGATGATTCCTTCAACACTGATGACAGAGATCACTGGAGCAAAAT 840
 Db 905 AAGACTTACCAAGATGATGATTCCTTCAACACTGATGACAGAGATCACTGGAGCAAAAT 964
 Qy 841 CCAGAAAGCTTGAATGCTGATGAGATGAGATGAGCACTGGAGATGTTGGCAGAGTGGC 900
 Db 965 CCAGAAAGCTTGAATGCTGATGAGATGAGATGAGCACTGGAGATGTTGGCAGAGTGGC 1024
 Qy 901 ATGATGATGAGCGGTGAGAGCTTTATATCAATGGAAGCACTCTCACTTTATTAAGATCTT 960
 Db 1025 ATGATGATGAGCGGTGAGAGCTTTATATCAATGGAAGCACTCTCACTTTATTAAGATCTT 1084
 Qy 961 GGGCTGCTTCAAAAGAGCACTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 1085 GGGCTGCTTCAAAAGAGCACTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1144
 Qy 1021 GTTGTGCTTCCAGTATTAATCAAGTAAAGTAAATTTTCCAACTACAGTCTGCTG 1080
 Db 1145 GTTGTGCTTCCAGTATTAATCAAGTAAAGTAAATTTTCCAACTACAGTCTGCTG 1204
 Qy 1081 ATGAGATCTGAGCGAGAACTTCTTACCACTGGGCTGCAATTAAGTGGAGTAAAG 1140
 Db 1205 ATGAGATCTGAGCGAGAACTTCTTACCACTGGGCTGCAATTAAGTGGAGTAAAG 1264
 Qy 1141 GCGAGGGCATCATGAGAGAGTAAAGAGCTGCTGAGAGCCCTCAATATCACTCAGGTC 1200
 Db 1265 GCGAGGGCATCATGAGAGAGTAAAGAGCTGCTGAGAGCCCTCAATATCACTCAGGTC 1324

Qy 1201 CTGAGCCATGAGAGAGGAGCAGACATCACTTTTGATCCAAAGTGGAGTCCGAGACC 1260
 Db 1325 CTGAGCCATGAGAGAGGAGCAGACATCACTTTTGATCCAAAGTGGAGTCCGAGACC 1384
 Qy 1261 AGTCTACTGATGATCTTAATCAAGATTTCTTTTCATCATCACTCCACGAGACACCATG 1320
 Db 1385 AGTCTACTGATGATCTTAATCAAGATTTCTTTTCATCATCACTCCACGAGACACCATG 1444
 Qy 1321 ACTGTCAATGATCCAAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTTGTTCTAT 1380
 Db 1445 ACTGTCAATGATCCAAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTTGTTCTAT 1504
 Qy 1381 GTTGTGACAGCATGAGAAAGATCTGCTAGTCC 1416
 Db 1505 GTTGTGACAGCATGAGAAAGATCTGCTAGTCC 1540

RESULT 4
 BC020689
 LOCUS
 DEFINITION Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418
 IMAGE:4251802, mRNA, complete cds.
 ACCESSION BC020689
 VERSION BC020689.1 GI:18088383
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Clontech Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
 Series: IRAL Plate: 35 Row: F Column: 13
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706386.

FEATURES

source

1. 1928
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /db_xref="taxon:9606"
 /clone="MGC:22418 IMAGE:4251802"
 /tissue_type="Skin, melanocytic melanoma, high MGR."
 /clone_lib="NIH_MGC_62"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 167. 1585
 /codon_start=1
 /product="plasma glutamate carboxypeptidase"
 /protein_id="AAH20689.1"
 /db_xref="GI:18088384"
 /translation="MKFLIPAFPGVHLISGSKAIKNGISKRTPEEIKETIASCG DVAKAINTLAVGAQNSYERLALIVDTVGPRLSGSKNEKAIQIWMYQNLQDGLK

CDS

167. 1585
 /codon_start=1
 /product="plasma glutamate carboxypeptidase"
 /protein_id="AAH20689.1"
 /db_xref="GI:18088384"
 /translation="MKFLIPAFPGVHLISGSKAIKNGISKRTPEEIKETIASCG DVAKAINTLAVGAQNSYERLALIVDTVGPRLSGSKNEKAIQIWMYQNLQDGLK

VHLPEVRI PHERGESBAMLEPRHIALGLGSSIGTPEGTAEVLVTSFDELQ
 RASERAGKIVVNOPYINYSRTVOYRTQAVENAKGALASLIRVASISYSPHTG
 IOEYDQVPIKIPACITVEDKEMMSRMAIGIKIOLKMGAKTVPYDSEFVYAEIT
 GSKYBEQVVLVSGHDSMDYGDMDGGGFIISWELSLIKDLGABKRTLLVMT
 AEROGVGAFOYVOLHAKVNI SNYSIWSBAGTLPFGLOPLOSERKARIMEVMSLT
 QPNTIYVLSHGSETDINFWIQAGVPSALIDLKYPFFHSHGDMTMDPKQNV
 AAAMVAVSVTVADMDEMLPRS"

BASE COUNT 554 a 435 c 466 g 473 t
 ORIGIN

Query Match 99.8%; Score 1412.8; DB 9; Length 1928;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAATTCCTTATCTGATTTTGGTGGTGTTCACCTTTATCCCTGCTGGG 60
 167 ATGAATTCCTTATCTGATTTTGGTGGTGTTCACCTTTATCCCTGCTGGG 226
 61 AAAGCTATATGCAAGAAATGCGATCTTAAGAGACTTTTGAAGAAATMAAGAAATA 120
 227 AAAGCTATATGCAAGAAATGCGATCTTAAGAGACTTTTGAAGAAATMAAGAAATA 286
 121 GCCAGCTGTGAGAGTGTGCTTAAGCAATCATCACTAGCTTTATGTTAAAGCCAG 180
 287 GCCAGCTGTGAGAGTGTGCTTAAGCAATCATCACTAGCTTTATGTTAAAGCCAG 346
 181 AACGATCTATGAGAGTGTGCTTCTGCTTATCTGTGAGCCAGCTAGTGGC 240
 347 AACGATCTATGAGAGTGTGCTTCTGCTTATCTGTGAGCCAGCTAGTGGC 406
 241 TCCAGAACTAGAAAAAGCCATCCAAATTATGTACCAAAACCTGCAGCAAGTGGCTG 300
 407 TCCAGAACTAGAAAAAGCCATCCAAATTATGTACCAAAACCTGCAGCAAGTGGCTG 466
 301 GAGAAATTTCACTGTGAGCCAGTGAATACCCCACTGGAGAGGGAGAAATCACT 360
 467 GAGAAATTTCACTGTGAGCCAGTGAATACCCCACTGGAGAGGGAGAAATCACT 526
 361 GTGATGCTGAGCCAAATTCATTAAGTATGCGATCCGGGCTTTGGAGAGCACTGGG 420
 527 GTGATGCTGAGCCAAATTCATTAAGTATGCGATCCGGGCTTTGGAGAGCACTGGG 586
 421 ACTCTCTCAGAAAGCAATTAAGCAAGTGTGCTGCTGCTCTTTCATGAACTGAG 480
 587 ACTCTCTCAGAAAGCAATTAAGCAAGTGTGCTGCTGCTCTTTCATGAACTGAG 646
 481 AGAAGGCTCTCAGAAAGCAAGTGTGCTGCTGCTCTTTCATGAACTGAG 540
 647 AGAAGGCTCTCAGAAAGCAAGTGTGCTGCTGCTCTTTCATGAACTGAG 706
 541 TCAAGAGCGGTGCAATTCAGAAAGCGGCGGTGGAAGCTGCCAGTGGGGCTTTG 600
 707 TCAAGAGCGGTGCAATTCAGAAAGCGGCGGTGGAAGCTGCCAGTGGGGCTTTG 766
 601 GCATCTCTCATTCGATCCGTGGCTCTCTTTCATCTCAAGTCTCTCAACAGTATTCAG 660
 767 GCATCTCTCATTCGATCCGTGGCTCTCTTTCATCTCAAGTCTCTCAACAGTATTCAG 826
 661 GAATATCAGAGATGCGTGGCCCAAAATTCAGACGCTGTATTAAGTGAAGATGCAAA 720
 827 GAATATCAGAGATGCGTGGCCCAAAATTCAGACGCTGTATTAAGTGAAGATGCAAA 886
 721 ATGATGTCAGAAATGCGTCTTCATAGGATCAAAATTTGTCATTCAGCTTAAGATGGGGCA 780
 887 ATGATGTCAGAAATGCGTCTTCATAGGATCAAAATTTGTCATTCAGCTTAAGATGGGGCA 946
 781 AAGACTACCAAGATGATGATCTCTTCAACACTGTAGAGAGATCACTGGAGCAAAATAT 840
 947 AAGACTACCAAGATGATGATCTCTTCAACACTGTAGAGAGATCACTGGAGCAAAATAT 1006
 841 CCAGAACAGGTTGTACTGTGATGAGCATCTGGAAGCTGGAGTGTGGCAGGGTGGC 900
 1007 CCAGAACAGGTTGTACTGTGATGAGCATCTGGAAGCTGGAGTGTGGCAGGGTGGC 1066

QY 901 ATGATGATGCGGTGAGCCCTTATATCATGGGAAGCACTCACTTATTAAGATCTT 960
 DB 1067 ATGATGATGCGGTGAGCCCTTATATCATGGGAAGCACTCACTTATTAAGATCTT 1126
 QY 961 GGGCTGCTCAGAAAGAGACTCTGCGGCTGCTCTGAGCTGAGAAAGAGAGTGA 1020
 DB 1127 GGGCTGCTCAGAAAGAGACTCTGCGGCTGCTCTGAGCTGAGAAAGAGAGTGA 1186
 QY 1021 GTTGGTCCCTTCAGATTAATCATGATACCAAGGTAATTTTCCAATCACTAGCTGGTG 1080
 DB 1187 GTTGGTCCCTTCAGATTAATCATGATACCAAGGTAATTTTCCAATCACTAGCTGGTG 1246
 QY 1081 ATGAGCTGACCGCAGGAACCTTCTTAACCACTGGGCTGCAATTCATCTGACAGTAAAG 1140
 DB 1247 ATGAGCTGACCGCAGGAACCTTCTTAACCACTGGGCTGCAATTCATCTGACAGTAAAG 1306
 QY 1141 GCCAGGGCATCATGAGAGGTTATGAGCTGCTGAGCCCTCAATATCACTAGCTG 1200
 DB 1307 GCCAGGGCATCATGAGAGGTTATGAGCTGCTGAGCCCTCAATATCACTAGCTG 1366
 QY 1201 CTGAGCCATGAGAAAGGAGACAGATCACTTTTGAATCCAACTGAGAGCTGAGCC 1260
 DB 1367 CTGAGCCATGAGAAAGGAGACAGATCACTTTTGAATCCAACTGAGAGCTGAGCC 1426
 QY 1261 AGTCTACTGATGACTTATACAGTATTTCTTCCATCACTCCACGAGACACCATG 1320
 DB 1427 AGTCTACTGATGACTTATACAGTATTTCTTCCATCACTCCACGAGACACCATG 1486
 QY 1321 ACTGTCAATGATCCAAAGCAGATGAATGTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1380
 DB 1487 ACTGTCAATGATCCAAAGCAGATGAATGTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1546
 QY 1381 GTTGTTCAGACATGAGAAAGATGCTGCTTACGTCC 1416
 DB 1547 GTTGTTCAGACATGAGAAAGATGCTGCTTACGTCC 1582

RESULT 5

AF107834 1472 bp mRNA linear PRI 12-JUL-1999
 LOCUS AF107834 Homo sapiens clone LCH1 aminopeptidase mRNA, complete cds.

DEFINITION AF107834.1 GI:5442029
 VERSION AF107834.1 GI:5442029
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1472)
 Liu, C.H., Lin, B.Y. and Chang, L.Y.
 Cloning of the human aminopeptidase gene

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1472)
 Liu, C.H., Lin, B.Y. and Chang, L.Y.
 Direct Submission

TITLE Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
 Sinica, RM 433, 128, Yen-Chun-Yuan Road SEC 2, Taipei 11529, Taiwan

JOURNAL Location/Qualifiers
 FEATURES 1. 1472
 source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="LCH1"
 /tissue_type="liver"
 /dev_stage="fetus"
 /note="secretory protein"
 /codon_start=1
 /product="aminopeptidase"
 /protein_id="A043214.1"
 /db_xref="GI:5442030"
 /translation="MKFLIPAPFGVHLISLCSGKAI CKNGISKRTFEEIKETIASCG

CDS

54..1472
 /note="secretory protein"
 /codon_start=1
 /product="aminopeptidase"
 /protein_id="A043214.1"
 /db_xref="GI:5442030"
 /translation="MKFLIPAPFGVHLISLCSGKAI CKNGISKRTFEEIKETIASCG

FEATURES
Source Location/Qualifiers
1..1860
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 515 a 423 c 451 g 471 t

ORIGIN

Query Match 99.4%; Score 1408; DB 6; Length 1860;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 ATGAATTCCTTATCTTGGATTTTGGTGGTTCACCTTTATCCCTGCTCTGG 60
129 ATGAATTCCTTATCTTGGATTTTGGTGGTTCACCTTTATCCCTGCTCTGG 188
61 AAACCTATATGCAAGAAATGGCACTCTAAGAGACTTTTGAAGAAATGAAGAAATA 120
189 AAACCTATATGCAAGAAATGGCACTCTAAGAGACTTTTGAAGAAATGAAGAAATA 248
121 GCCAGCTGTGAGATGTTGTAAAGCAATCACTAGCTGTTTATGTTAAAGCCAG 180
249 GCCAGCTGTGAGATGTTGTAAAGCAATCACTAGCTGTTTATGTTAAAGCCAG 308
181 AACAGATCTATGAGCGATTTGCACTTCTGTTGATCTGTTGACCCAGACTGAGTGC 240
309 AACAGATCTATGAGCGATTTGCACTTCTGTTGATCTGTTGACCCAGACTGAGTGC 368
241 TCCAGAACTTAAAGAAAGCCATCCAAATTTATGCCAAACCTGACAGCAAGATGGCTG 300
369 TCCAGAACTTAAAGAAAGCCATCCAAATTTATGCCAAACCTGACAGCAAGATGGCTG 428
301 GAGAAAGTTTCACTGAGACGAGTGAATACCCCACTGGAGAGAGAGAGAAATCACT 360
429 GAGAAAGTTTCACTGAGACGAGTGAATACCCCACTGGAGAGAGAGAAATCACT 488
361 GTGATGCTGAGCCAGAAATTCATAGATAGCCATCTCGGTCTTGGACGACGATTTGG 420
489 GTGATGCTGAGCCAGAAATTCATAGATAGCCATCTCGGTCTTGGACGACGATTTGG 548
421 ACTCTCTCAGAAAGCATTACGACAGAGTTCTGTTGATCTCTTTGATGAATCTGAG 480
549 ACTCTCTCAGAAAGCATTACGACAGAGTTCTGTTGATCTCTTTGATGAATCTGAG 608
481 AGAAGGGCTCAGAGAAAGGAGGAAATGTTGTTTAAACCAACTTAACTAATCACTAC 540
609 AGAAGGGCTCAGAGAAAGGAGGAAATGTTGTTTAAACCAACTTAACTAATCACTAC 668
541 TCAAGACGCTGCAATACCAAGCCAGGGGCGGTGGAAGCTGCCAAGGTGGGGCTTTG 600
669 TCAAGACGCTGCAATACCAAGCCAGGGGCGGTGGAAGCTGCCAAGGTGGGGCTTTG 728
601 GCATCTCTCATTCGATCCGTTGCTCTTCTTCATCTAAGATCTCTCACAGGATTTAG 660
729 GCATCTCTCATTCGATCCGTTGCTCTTCTTCATCTAAGATCTCTCACAGGATTTAG 788
661 GAATACAGAGATGGGTCGCCAAATTCACACAGCTGTATTAAGGTGAAGATGCGAA 720
789 GAATACAGAGATGGGTCGCCAAATTCACACAGCTGTATTAAGGTGAAGATGCGAA 848
721 ATGATGTCAGAAATGCTTCTCATGAGATCAAAATTTGCTATGACTTAAAGTGGGGCA 780
849 ATGATGTCAGAAATGCTTCTCATGAGATCAAAATTTGCTATGACTTAAAGTGGGGCA 908
781 AAGACCTAACCAAGATCTGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATAT 840
909 AAGACCTAACCAAGATCTGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATAT 968
841 CCAGAAACAGGTTTACTGCTGAGTGAAGATCTGACACCTGGAGATGTTGGCAGAGTGGC 900
969 CCAGAAACAGGTTTACTGCTGAGTGAAGATCTGACACCTGGAGATGTTGGCAGAGTGGC 1028
901 ATGATGATGGCGGTGAGCTTTATATCATGGAAGACATCTCACTTAAAGATCTT 960

1029 ATGATGATGGCGGTGAGCTTTATATCATGGAGACATCTCACTTAAAGATCTT 1088
961 GCGCTGCGTCCAAAGAGACATCTGCGGTGCTCTGAGATCTGCAAGAAAGAGTGA 1020
1089 GGGCTGCGTCCAAAGAGACATCTGCGGTGCTCTGAGATCTGCAAGAAAGAGTGA 1148
1021 GTTGGTGCCTCCAGATATATCATGTTACAAAGGTAATATTTCACTACAGTGGTG 1080
1149 GTTGGTGCCTCCAGATATATCATGTTACAAAGGTAATATTTCACTACAGTGGTG 1208
1081 ATGAGTCTGACGAGAAACCTTCTTACCACTGGCTGCAATTCATCTGAGCAAAAG 1140
1209 ATGAGTCTGACGAGAAACCTTCTTACCACTGGCTGCAATTCATCTGAGCAAAAG 1268
1141 GCCAGGCGCATGAGAGAGTTATGAGCTGCTGCAAGCCCTCAATATCACTCAGTGC 1200
1269 GCCAGGCGCATGAGAGAGTTATGAGCTGCTGCAAGCCCTCAATATCACTCAGTGC 1328
1201 CTGAGCCATGAGAAAGGAGACATCACTTTGATCCAAAGCTGAGTGCCTGGAGCC 1260
1329 CTGAGCCATGAGAAAGGAGACATCACTTTGATCCAAAGCTGAGTGCCTGGAGCC 1388
1261 AGTCTACTGATGACTTATACAAATTTCTTCTTCATCACTGCCAGAGACCATATG 1320
1389 AGTCTACTGATGACTTATACAAATTTCTTCTTCATCACTGCCAGAGACCATATG 1448
1321 ACTGTCATGATCCAAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTTTCTTAT 1380
1449 ACTGTCATGATCCAAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTTTCTTAT 1508
1381 GTTGTGACAGCATGAGAAAGTCTGCTGAGTCC 1416
1509 GTTGTGACAGCATGAGAAAGTCTGCTGAGTCC 1544

RESULT 7
AK075132
LOCUS
DEFINITION
Homo sapiens cDNA FLJ90651 fis, clone PLACE100482, moderately similar to Rattus norvegicus hematopoietic lineage switch 2 related protein (Hlzf-xp) mRNA.
AK075132
ACCESSION
AK075132.1 GI:22761022
VERSION
Oligo capping, fis (full insert sequence).
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Isegai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Kojima,S., Nagahara,K., Masuno,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Ninomiya,K.
NEO human cDNA sequencing project
TITLE
NEO human cDNA sequencing project
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1860)
Isegai,T. and Otsuki,T.
AUTHORS
DIRECT SUBMISSION
JOURNAL
Submitted (25-MAR-2003) Takao Isegai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT
NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).
FEATURES
source
1..1860
Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004482"
/issue_type="Placenta"
/clone_id="PLACE1"
/notes="Cloning vector: PME18SFL3"
129.1547
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC11423.1"
/db_xref="GI:22761023"
/translation="MKFLIFAFPGVHLISLCSKALCKNGISKREIEIEBIAISG
DVAKIIMLAYGKONRSYERLALVDVTGPRLSGKLEKAIQIMYONQOQLEK
VHLEPVRIPHEWGESAVMLEPRHKIAIIGLSSICTPPEGTAIVLVTSFHDG
RASEBARKIIVYNOPYINVSRYOYRTOGVEAKKVALASILRSVASPSIYSPHTG
IOERODGPKIPIKICITYEDBEMSRMSHGIKIVIOUKMAKATYPTDTSNTVAET
GSKYPEOVVALSHLSDVDGQAMDGGAPISWEALSLIKDGLRKRRLRLVLMWT
AEBOGVGAFQYQOLHKVNIISNYSIWMBSDGTPLPTGLQFTGSEKAKAIVBEYWSLL
OPLNTVOYLSHGEETDINFMIQAGVPGASLDDLYKYFFPHSHQDTWTVDPKOMNY
AAAVWAVSYVVADEMLPRS"
BASE COUNT      515 a      423 c      451 g      471 t
ORIGIN
Query Match      99.4%; Score 1408; DB 9; Length 1860;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 ATGAATTCCTTATCTTGCAATTTTCGCTGCTGTCACCTTTATCCCTGCTCTGGG 60
DB 129 ATGAATTCCTTATCTTGCAATTTTCGCTGCTGTCACCTTTATCCCTGCTCTGGG 188
OY 61 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAAATA 120
DB 189 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAAATA 248
OY 121 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 180
DB 249 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 308
OY 181 AACAGATCTTATGACGATGGCACTTCTGTTGATCTGTTGACCCAGACTGAGTGGC 240
DB 309 AACAGATCTTATGACGATGGCACTTCTGTTGATCTGTTGACCCAGACTGAGTGGC 368
OY 241 TCCAGAACCTAAGAAAGCAATCCAAATTAATGTAACCAAACTGACCAAGATGGGCTG 300
DB 369 TCCAGAACCTAAGAAAGCAATCCAAATTAATGTAACCAAACTGACCAAGATGGGCTG 428
OY 301 GAGAAAGTTCACTGTGAGCCAGTGAATATCCCACTGGAGAGGGAGAAATCAGCT 360
DB 429 GAGAAAGTTCACTGTGAGCCAGTGAATATCCCACTGGAGAGGGAGAAATCAGCT 488
OY 361 GTGATGCTGAGCCAAAGATTCATTAAGATGACCATCTGGGCTTTGGCAGCAGCATTTGGG 420
DB 489 GTGATGCTGAGCCAAAGATTCATTAAGATGACCATCTGGGCTTTGGCAGCAGCATTTGGG 548
OY 421 ACTGCTCAGAAAGGATTAAGCAGAAAGTCTGGGAGTGAACCTCTTTGATGAATACTGAG 480
DB 549 ACTGCTCAGAAAGGATTAAGCAGAAAGTCTGGGAGTGAACCTCTTTGATGAATACTGAG 608
OY 481 AGAAGGGGCTCAGAAAGAGGAGATTTGTTTATTAACCAACTTATACATCACTAC 540
DB 609 AGAAGGGGCTCAGAAAGAGGAGATTTGTTTATTAACCAACTTATACATCACTAC 668
OY 541 TCAAGACGCTGCAATACCGAAGCGAGGGGCGGTGAAGCTGCCAGGTGGGGCTTTG 600
DB 669 TCAAGACGCTGCAATACCGAAGCGAGGGGCGGTGAAGCTGCCAGGTGGGGCTTTG 728
OY 601 GCATCTCTCATTCGATCCGTGGCTCTTCTTCATCTACAGTCTCTCAACAGATATTAG 660
DB 729 GCATCTCTCATTCGATCCGTGGCTCTTCTTCATCTACAGTCTCTCAACAGATATTAG 788
OY 661 GAATATCAGATATGGGTGCCCAAAATTCACACAGCCTGTATTAAGGTGAAGATGCAGAA 720

```

```

|||||
DB 789 GAATATCAGATATGGGTGCCCAAAATTCACACAGCCTGTATTAAGGTGAAGATGCAGAA 848
OY 721 ATGATGTCAGAAATAGGCTTTCTCATAGGATCAAAATTTGCTATTGCTAAAGATGGGGCA 780
DB 849 ATGATGTCAGAAATAGGCTTTCTCATAGGATCAAAATTTGCTATTGCTAAAGATGGGGCA 908
OY 781 AAGACTTACCAGATATCTGATATCTTCAACACTGTACAGAGATCACTGGGAGCAATAT 840
DB 909 AAGACTTACCAGATATCTGATATCTTCAACACTGTACAGAGATCACTGGGAGCAATAT 968
OY 841 CCAGAACAGGTTGACTGTGTCAGTGAACATCTGACAGCTGGAGATGTTGGCAGGTCGCC 900
DB 969 CCAGAACAGGTTGACTGTGTCAGTGAACATCTGACAGCTGGAGATGTTGGCAGGTCGCC 1028
OY 901 ATGATATATGGCGGTGAGAGCTTTATATCATATGGAACACCTTCATTTAAGATCTT 960
DB 1029 ATGATATATGGCGGTGAGAGCTTTATATCATATGGAACACCTTCATTTAAGATCTT 1088
OY 961 GGGCTGCGTCCAAAGAGAGACTCTGCGGCTGCTGTGACTGACAGAGAAACAAGTGA 1020
DB 1089 GGGCTGCGTCCAAAGAGAGACTCTGCGGCTGCTGTGACTGACAGAGAAACAAGTGA 1148
OY 1021 GTTGTGCTCTTCCAGATATTAATCACTTAACAGGTAATAATTTCAACTACATCTGCTG 1080
DB 1149 GTTGTGCTCTTCCAGATATTAATCACTTAACAGGTAATAATTTCAACTACATCTGCTG 1208
OY 1081 ATGAGATCTGACGAGAGAACCTTTTACCACTGAGGCTGCAATTCATCTGGAGGAAAG 1140
DB 1209 ATGAGATCTGACGAGAGAACCTTTTACCACTGAGGCTGCAATTCATCTGGAGGAAAG 1268
OY 1141 GCCAGGGCCATCATGAGAGAGGATTATGAGCTGTGACGCCCTCAATATCATCTCAGGTC 1200
DB 1269 GCCAGGGCCATCATGAGAGAGGATTATGAGCTGTGACGCCCTCAATATCATCTCAGGTC 1328
OY 1201 CTGAGCCATGAGAGAGGAGACAGACATCACTTTTGAATTCAGCTGAGTGCCTGAGCC 1260
DB 1329 CTGAGCCATGAGAGAGGAGACAGACATCACTTTTGAATTCAGCTGAGTGCCTGAGCC 1388
OY 1261 AGTTACTTATGATGATTAATACAGATATTTCTTTCATCATCTCCACGAGACACCATG 1320
DB 1389 AGTTACTTATGATGATTAATACAGATATTTCTTTCATCATCTCCACGAGACACCATG 1448
OY 1321 ACTGTCAATGATCCAAAGAGAGATGATGTTGCTGCTGTGTTGGGCTGTTCTTAT 1380
DB 1449 ACTGTCAATGATCCAAAGAGAGATGATGTTGCTGCTGTGTTGGGCTGTTCTTAT 1508
OY 1381 GTTGTGCAAGCATGAGAAAGAAATGCTGCTAGTGC 1416
DB 1509 GTTGTGCAAGCATGAGAAAGAAATGCTGCTAGTGC 1544

RESULT 8
AF119386 1794 bp mRNA linear PRI 20-MAY-1999
LOCUS Homo sapiens blood plasma glutamate carboxypeptidase precursor
DEFINITION (RCCP) mRNA, complete cds.
VERSION AF119386
KEYWORDS AF119386.1 GI:4877697
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1794)
AUTHORS Gingras, R., Richard, C., Bl-Alfy, M., Morales, C.R., Potier, M. and
Pehzetsky, A.V.
Purification, cDNA cloning, and expression of a new human blood
plasma glutamate carboxypeptidase homologous to
N-acetyl-aspartyl-alpha-glutamate
carboxypeptidase/prostate-specific membrane antigen
JOURNAL J. Biol. Chem. 274 (17), 11742-11750 (1999)
MEDLINE 99223495

```

PUBMED 10206990
REFERENCE 2 (bases 1 to 1794)
AUTHORS Gingras, R., Richard, C., El-Alfy, M., Morales, C. R., Poirier, M. and Pehezhetsky, A. V.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1999) Medical Genetics, Sainte-Justine Hospital, Montreal University, 3175 Cote Sainte-Catherine, Montreal, QU H3T 1C5, Canada

FEATURES
source
1. 1794
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q22.2; between DBS257 and DBS270"
1. 1794
/gene="PGCP"
25. 1650
/EC number="3.4.17.21"
/function="endopeptidase"
/note="member of M28 peptidase family; similar to glutamate carboxypeptidase II and to N-acetyl-aspartyl-alpha-glutamate carboxypeptidase (MALADase); similar to sequences with GenBank Accession Numbers R18560 and G07404; prostate-specific membrane antigen (PSMA)"
/codon_start=1
/product="blood plasma glutamate carboxypeptidase precursor"
/protein_id="AAD31438.1"
/db_xref="GI:4877698"
/translation="MKFLIFAFGFGVHLISGSKAI CKNGISKRTFEEIKETIASG DVAKILINLVYGAQKRSYERIALVDYDTPRSGSKNLEKAIQIYVOMI OODGLK VHLPEVPIPMERGESAVMLERIKHILAILGLSSGTPEGGITAEVLVYTSDEIQ RRASEARKIVVINOPIYINRSIVQYRTQGAENAKVAGALASLIRVASPSIYSPHG IOEVQDVPRIPTACITVEDAEKMSRASHQIKIVLQKGAQKYPIDSPNTVAETG GSKYREQVVLVSGHDSMDVQCGAMDGGGAFISWEALSLIKDLGLPKETRLVLT ABEQGVGAFOYQHLHKVNI SNVSLVMSDSAGTLPGLDPTGSEKRALIMEEWSLL OPTINIOVLSHGEGTDINPMIOAGVPGASLIDDLKKYFFPHSHGDMTWDPSRWML LILFGLFLMLLOTWKCKCCUQPRNSKKEKTSCTPFGILGILQMLPHITTSNSSK HNSISCFLLSLFSLKPSDSRKNHSPLEPTT"
25. 96
/gene="PGCP"
157. 1647
/gene="PGCP"
/product="blood plasma glutamate carboxypeptidase"
205. 207
/gene="PGCP"
/note="putative; glycosylation site"
559. 561
/gene="PGCP"
/note="putative; glycosylation site"
1081. 1083
/gene="PGCP"
/note="putative; glycosylation site"
1090. 1092
/gene="PGCP"
/note="putative; glycosylation site"
1210. 1212
/gene="PGCP"
/note="putative; glycosylation site"
/note="putative; glycosylation site"
BASE COUNT 522 a 391 c 422 g 457 t 2 others
ORIGIN

Query Match 98.8%; Score 1399.2; DB 9; Length 1794;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1412; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 ATGAATTCCTTATCTTCGATTTTTCGATGCTGCTTATACCTTATCCTGCTCTGGG 60
DB 25 ATGAATTCCTTATCTTCGATTTTTCGATGCTGCTTATACCTTATCCTGCTCTGGG 84

QY 61 AAAGCTATATGACAGAAATGSCATCTCTAAGAGGACTTTTGAAGAAATTAAGAGAAATA 120
DB 85 AAAGCTATATGACAGAAATGSCATCTCTAAGAGGACTTTTGAAGAAATTAAGAGAAATA 144
QY 121 GCCAGCTGTGAGAGTGTGCTTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 180
DB 145 GCCAGCTGTGAGAGTGTGCTTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 204
QY 181 AACGATCTATGAGGATGAGGACTTCTGCTTGA TACTGTGAGACCAGACTGAGTGGC 240
DB 205 AACGATCTATGAGGATGAGGACTTCTGCTTGA TACTGTGAGACCAGACTGAGTGGC 264
QY 241 TCAGAACTAG 300
DB 265 TCAGAACTAG 324
QY 301 GAGAAAGTTCACTGAG 360
DB 325 GAGAAAGTTCACTGAG 384
QY 361 GTGATGCTGAG 420
DB 385 GTGATGCTGAG 444
QY 421 ACTCTTCAG 480
DB 445 ACTCTTCAG 504
QY 481 AGAAGGCTTCAG 540
DB 505 AGAAGGCTTCAG 564
QY 541 TCAAG 600
DB 565 TCAAG 624
QY 601 GCACTCTCATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 660
DB 625 GCACTCTCATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 684
QY 661 GAATACAGAGATGCGTGGCCCAAAATTCGACAGCTGTATTA CGGTGAAGATGCAGAA 720
DB 685 GAATACAGAGATGCGTGGCCCAAAATTCGACAGCTGTATTA CGGTGAAGATGCAGAA 744
QY 721 ATGATGCAAGATGCGTGGCCCAAAATTCGACAGCTGTATTA CGGTGAAGATGCAGAA 780
DB 745 ATGATGCAAGATGCGTGGCCCAAAATTCGACAGCTGTATTA CGGTGAAGATGCAGAA 804
QY 781 AAGACTACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGAGCAATAT 840
DB 805 AAGACTACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGAGCAATAT 864
QY 841 CCAAGACAGGTTGATCTGATCACTGATCACTGATCACTGATCACTGATCACTGATCACT 900
DB 865 CCAAGACAGGTTGATCTGATCACTGATCACTGATCACTGATCACTGATCACTGATCACT 924
QY 901 ATGATGATGCGGATGAGAGCTTATATCATGGAAGACATCTCACTATTATTAAGATCT 960
DB 925 ATGATGATGCGGATGAGAGCTTATATCATGGAAGACATCTCACTATTATTAAGATCT 984
QY 961 GGGCTGCTGCAAG 1020
DB 985 GGGCTGCTGCAAG 1044
QY 1021 GTTGTGCTTCCAGAT 1080
DB 1045 GTTGTGCTTCCAGAT 1104
QY 1081 ATGAGATGATGAG 1140
DB 1105 ATGAGATGATGAG 1164
QY 1141 GCCAGGCGCATGATGAG 1200


```

Db      1165 GCCAGGGCCATCATGAGAGAGGTATGAGCCCTGCTGACAGCCCTCAATATGACTCAGGTC 1224
Qy      1201 CTGAGCCATGAGAGAGGACAGACACTCACTTTTGGATCCAGCTGAGTGCTGGAGACC 1260
Db      1225 CTGAGCCATGAGAGAGGACAGACACTCACTTTTGGATCCAGCTGAGTGCTGGAGACC 1284
Qy      1261 AGTCACTTGAATGACTTATACAGATTTCTTCTCCATCACTCCACGAGAGACACATG 1320
Db      1285 AGTCACTTGAATGACTTATACAGATTTCTTCTCCATCACTCCACGAGAGACACATG 1344
Qy      1321 ACTGTCATGATCCAAAGACAGATGATGTGCTGCTGTTGGCTGTTGTTCTTAT 1380
Db      1345 ACTGTCATGATCC-CACAGATGATGATGTGCTGCTGTTGGCTGTTGTTCTTAT 1403
Qy      1381 GTTGTGACAGATGAGAGAAATGCTGCTAGGTC 1416
Db      1404 GTTGTGACAGATGAGAGAAATGCTGCTAGGTC 1439

RESULT 9
BC037067 1796 bp mRNA linear ROD 16-APR-2003
LOCUS     Mus musculus plasma glutamate carboxypeptidase, mRNA (cDNA clone
DEFINITION
MGC:46857 IMAGE:4989536), complete cds.
ACCESSION BC037067
VERSION   BC037067.1 GI:22477497
KEYWORDS  MGC.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1796)
AUTHORS   Strausberg,R.L., Feingold,F.A., Grouse,L.H., Derge,J.G.,
          Klusner,R.D., Collins,F.S., Wagner,J., Shenmen,C.M., Schuler,G.D.,
          Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
          Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
          Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
          Stjepicon,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
          Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Topolynski,S.,
          Cantinini,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
          Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
          McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
          Wooley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Holys,S.W.,
          Viallano,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
          Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
          Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
          Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
          Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
          Butlerfield,Y.S., Krzywinski,M.I., Skalka,U., Smellus,D.E.,
          Schermer,A., Schein,J.E., Jones,S.J. and Marra,M.A.,
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL   2388257
MEDLINE   12477932
PUBMED    2 (bases 1 to 1796)
REFERENCE Strausberg,R.
TITLE     Direct Submission
AUTHORS   Submitted (23-AUG-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
          Contact: MGC help desk
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC),
          Gaithersburg, Maryland!
          Web site: http://www.nisc.nih.gov/

```

Contact: nisc.mgc@nih.gov
 Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
 Maduro,Q.L., Maelillo,C., Maskeri,B., Mastrilli,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,R., Stancirpop,S., Thomas,P.J., Touchman,J.W.,
 Tsurgouon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAX Plate: 81 Row: 9 Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 9055233.

FEATURES
 source
 location/Qualifiers
 1..1796
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:46857 IMAGE:4989536"
 /issue_type="Colon, normal, 5 month old male mouse."
 /clone_id="NCI CGAP_Co24"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1..1796
 /gene="Pscp-pending"
 /note="synonym: H152, 1190003P12Rik"
 /db_xref="locusid:54381"
 /db_xref="mgi:1889205"
 63..1475
 /codon_start=1
 /product="plasma glutamate carboxypeptidase"
 /protein_id="AAH37067.1"
 /db_xref="GI:22477498"
 /db_xref="locusid:54381"
 /translaton="MRSLEFLPLVILALSGKAVPKNGVORTPREKEIKENYEDV
 AKTINAVYGVKQNSYERLGLVLDLHLSGKAVPKNGVORTPREKEIKENYEDV
 LEQYRIPMERGESAVMLEPRIRHKAIIILGSGISGTPREGITRAVLVASFPELQQR
 ASEBGRKIVYNQPYGTEKTVYRQVGAIVKAVASLIQSVASFYSFPTGIG
 KYDQVGEPIPTACITVEDAMBSMRKRIYIHLEMGAKYEPDVSFNTVAETGS
 MYBEVAVLVGHDSMDVQGLDGGAFISWELSLVLDLGLRPRTIRLVLTAE
 EOGGISOYEYEHKANIISKYSLMEADSGPLPTGLQFTGSDKARIMKENVMLQP
 LNTKPSNGEGDINFIQIAGVPGASLRDLYKYPFFHSHDVTVMQKNNVAA
 AWAVAIVAVVADMDHLPRS"

BASE COUNT 526 a 372 c 418 g 480 t
 ORIGIN

Query Match 78.9%; Score 1116.8; DB 10; Length 1796;
 Best Local Similarity 87.9%; Pred. No. 3.1e-108;
 Matches 1211; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```

Qy      33 TGTTCACCTTTTATCCCTGTGCTCTGGAAAGCTATATGACAGATGCACTCTTAAGAG 92
Db      89 TGTTCACCTTTTATCCCTGTGCTCTGGAAAGCTATATGACAGATGCTTTCTGCGC 148
Qy      93 GACTTTGAAGAATATAAAGAAGAAATAGCCAGCTGTGAGATGTTCTTAAGAAATCAT 152
Db      149 AACATTTTCGAGAAATATAAAGAAGAAATAGCCAGCTATGAAGATGTTCTTAAGAAATAT 208
Qy      153 CAACCTAGCTGTTTATGATTAAGCCGAGAACAGATCTTATGAGGATGTCGACTTCTGCT 212
Db      209 CAACCTAGCTGTTTATGATTAAGCCGAGAACCGGTCTTATGAGGCTTTGGACTCTTAGT 268
Qy      213 TGATACCTGTGAGCCAGACTGAGTGCTCCAGAACCTTAAGAAAGCCATCCAAATAT 272
Db      269 TGATACCTGTGAGCCAGACTGAGTGCTCCAGAACCTTAAGAAAGCCATCCAAATAT 328
Qy      273 GTACCAAAACCTGACGAGAAAGTGCGCTGAGAAAGTTCACTGAGGCAAGTGAATATCC 332
Db      329 GTACCAAAACCTGACGAGAAAGTGCGCTGAGAAAGTTCACTGAGGCAAGTGAATATCC 388

```



```

Qy 333 CCACCTGGAGAGGGGAGAGAGATCAGCTGTGATCTGAGCCAGAAATTCATAGATAGC 392
Db 389 CCACTGGAGAGGGGAGAGAGATCTGAGTATCTTGAGCTTGAAATTCAGAGATGGC 448
Qy 393 CATCTGGGCTTGGGAGAGATTTGGGACTCTCCAGAAAGGCTTACAGAGAGTTCT 452
Db 449 TATTTAGGCTTGGGAGAGATTTGGGACTCTCCAGAGAGGCTTACAGAGAGTTCT 508
Qy 453 GGTGTGACCTCTTTCAGTGAATGACAGAGAGGGGCTTCAGAACCAAGAGGAGATTGT 512
Db 509 GGTGTGACCTCTTTCAGTGAATGACAGAGAGGGGCTTCAGAACCAAGAGGAGATTGT 568
Qy 513 TGTATTACCAACCTTACATCACTACTCAAGAGCGGTGCAATACCGAACGAGGGAGC 572
Db 569 TGTATTACCAACCTTACATCACTACTCAAGAGCGGTGCAATACCGAACGAGGGAGC 628
Qy 573 GGTGGAAGCTGCGCAAGGTGGGGCTTTGGGACTCTCTCATTTGATTCGCTGCTCTTCTC 632
Db 629 TGTGGAAGCTGCGCAAGGTGGGGCTTTGGGACTCTCTCATTTGATTCGCTGCTCTTCTC 688
Qy 633 CATCTACGCTCCACACAGGATTTCAAGAAATACAGAGTGGCGCCAAATTCACAC 692
Db 689 CATCTACGCTCCACACAGGATTTCAAGAAATACAGAGTGGCGCCAAATTCACAC 748
Qy 693 AGCCTGATTACCGGTGGAAGATGAGAAATGATGTCAAGATGGCTTCTCATGGGATCAA 752
Db 749 AGCCTGATTACCGGTGGAAGATGAGAAATGATGTCTCGAATGGCTTCTCATGGGATCAA 808
Qy 753 AATTGTCATTACGCTAAAGATGGGGGCAAGACCTTACCCAGATATCTGATTCCTTCAA 812
Db 809 AATTGTCATTACGCTAAAGATGGGGGCAAGACCTTACCCAGATATCTGATTCCTTCAA 868
Qy 813 TGTAGAGAGATCACTGGGAGCAATATTCAGAACAGTGTGATGTGATGATGAGATCT 872
Db 869 TGTAGAGAGATCACTGGGAGCAATATTCAGAACAGTGTGATGTGATGATGAGATCT 928
Qy 873 GGAAGCTGGGATTTGGGAGCGGTGCAATGATGATGATGATGATGATGATGATGATG 932
Db 929 GGAAGCTGGGATTTGGGAGCGGTGCAATGATGATGATGATGATGATGATGATGATG 988
Qy 933 GGAAGCTGCTCACTTATTAAGATCTTGGGCTGCTGCTCAAAAGAGACTGCGGCTGGT 992
Db 989 GGAAGCTGCTCACTTATTAAGATCTTGGGCTGCTGCTCAAAAGAGACTGCGGCTGGT 1048
Qy 993 GCTCTGAGCTGAGAGAAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATG 1052
Db 1049 GCTCTGAGCTGAGAGAAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATG 1108
Qy 1053 GGTAAATATTTTCAACTACAGTCTGATGATGATGATGATGATGATGATGATGATGATG 1112
Db 1109 GGTAAATATTTTCAACTACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1168
Qy 1113 TGGGCTGCAATTCATCGGAGTGAAGAGCCAGGGGCAATCATGAGAGGTTTATGAGCCT 1172
Db 1169 TGGGCTGCAATTCATCGGAGTGAAGAGCCAGGGGCAATCATGAGAGGTTTATGAGCCT 1228
Qy 1173 GCTGAGACCCCTCAATCATCAAGTCTGAGGCTGAGGCTGAGAGAGAGAGAGAGAGAG 1232
Db 1229 TCTGCAACCCCTCAATCATCAAGTCTTATGATGATGATGATGATGATGATGATGATGATG 1288
Qy 1233 TTGGATCAAGAGTGAAGTGGCTGAGGCAAGTCTGATGATGATGATGATGATGATGATGATG 1292
Db 1289 TTGGATCAAGAGTGAAGTGGCTGAGGCAAGTCTGATGATGATGATGATGATGATGATGATGATG 1348
Qy 1293 CTTCATCACTCCACGAGAGACACATGATGATGATGATGATGATGATGATGATGATGATG 1352
Db 1349 TTTCATCACTCCACGAGAGACACATGATGATGATGATGATGATGATGATGATGATGATG 1408
Qy 1353 TGTGCTGTTTGGGCTGTTGTTTCTTATGTTTGGAGAGATGAGAGAGAGAGAGAGAGAG 1412
Db 1409 TGTGCTGTTTGGGCTGTTGTTTCTTATGTTTGGAGAGATGAGAGAGAGAGAGAGAGAG 1468
Qy 1413 GTCC 1416

```

```

Db 1469 GTCC 1472
RESULT 10
AF107835
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1596
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="LCH-M1"
/tissue_type="liver"
184..1596
/note="secretory protein"
/codon_start=1
/product="aminopeptidase"
/protein_id="A043215.1"
/db_xref="GI:5442032"
/translation="MRSLEFLFVHLAAGSKAVPNKGVDSORPREIKELANEDV
AKKINLNAVYKQNSRYERLGLVDIVTERLGSKRLEKAIQIMONLOODGLENH
LEQVRIPIHMERGESAVMLEPRIRHNAIILGSSIGPREGITAEVLVAVSPBLQNR
ASBARKIIVNOPYKTEKTVYRQVGAFAKAVAVASLIQSVASFYSFPTGIGI
KYODGVPKIPFACITVEDAMSMRMSRGRKIYIHEMGAKTYPDSDPNTVAEITGS
MYEEVILAVSGHLDSDVDGQALDDGGARISMEALSLVVDLGLRPRKTRLVMTWE
EOGGIGASOYVEHLKXANISKYSLVMEADSGTLPDGLQFGSDKARAIKEXVNILOP
LNTKTFPSNBEGRDINFWIOAGVPGASLRDLKYKFFHSHSDTMTVMDPKQNTVA
AVMAVVAAYVVDMDMLPRS"
BASE COUNT 415 a 357 c 421 g 403 t
ORIGIN
Query Match 78.8%; Score 1115.2; DB 10; Length 1596;
Best Local Similarity 87.9%; Pred. No. 8.8e-308;
Matches 1216; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
Qy 33 TGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATCTTAAGAG 92
Db 210 TGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATCTTAAGAG 269
Qy 93 GACTTTGAAGAAATTAAGAAAGATAGCCAGCTGTGAGATGTTGCTTAAGCAATCAT 152
Db 270 AACATTTGAGAAATTAAGAAAGATAGCCAGCTGTGAGATGTTGCTTAAGCAATCAT 329
Qy 153 CAACCTGCTGTTTATGTGTAAGCCCGAGACAGATCCTATGAGGATGTTGCACTTCTGCT 212
Db 330 CAACCTGCTGTTTATGTGTAAGCCCGAGACAGATCCTATGAGGATGTTGCACTTCTGCT 389
Qy 213 TGAATCTGTGAGCCGAGCTGAGTGGCTCCAGAAACCTTGAAGAAACCTTGAATTTAT 272
Db 390 TGAATCTGTGAGCCGAGCTGAGTGGCTCTTAAGAACTTGAAGAAACCTTGAATTTAT 449
Qy 273 GTACCAAAACCTGAGAGATGAGGCTGAGAAAGTTCACTGAGGACAGTGAAGATACC 332
Db 450 GTACCAAAACCTGAGAGATGAGGCTGAGAAAGTTCACTGAGGACAGTGAAGATACC 509

```

```

Qy 333 CCACGTGGAGAGAGAGAGAGATCAGCTGTGATGCTGGAGCCAGAAATTCATTAAGATAGC 392
Db 510 CCACTGGAGAGAGAGAGAGAGATCGAGTGTGATGCTGGAGCCAGAAATTCATTAAGATAGC 569
Qy 393 CATCTGGGTCTTGGAGAGAGATTTGGAGCTTCCAGAGAGGCTTACAGAGAAATTTCT 452
Db 570 TATCTAGGTCTTGGAGAGAGATTTGGAGCTTCCAGAGAGGCTTACAGAGAAATTTCT 629
Qy 453 GGTGTGAGCTCTTTCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 512
Db 630 GGTGTGAGCTCTTTCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689
Qy 513 TGTATTAAACCACTTACATCACTACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 572
Db 690 TGTATTAAACCACTTACATCACTACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 749
Qy 573 GGTGTGAGCTCTTTCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 632
Db 750 TGTGTGAGCTCTTTCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809
Qy 633 CATCTAGAGCTCTCAACAGAGATTTACAGAGATTTACAGAGATTTACAGAGATTTACAGAG 692
Db 810 CATCTAGAGCTCTCAACAGAGATTTACAGAGATTTACAGAGATTTACAGAGATTTACAGAG 869
Qy 693 AGCTGTATTACGGTGAAGATGAGAGAGATGATCAAGAGATGATCAAGAGATGATCAAGAG 752
Db 870 AGCTGTATTACGGTGAAGATGAGAGAGATGATCAAGAGATGATCAAGAGATGATCAAGAG 929
Qy 753 AATGTCTATTCAGTAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
Db 930 AATGTCTATTCAGTAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 989
Qy 813 TGTGTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 872
Db 990 TGTGTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1049
Qy 873 GAGAGAGCTGGAGATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 932
Db 1050 GAGAGAGCTGGAGATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1109
Qy 933 GGAAGAGCTCTCACTTATTAAAGATCTTGGAGCTGCTCAAGAGAGAGAGAGAGAGAGAGAG 992
Db 1110 GGAAGAGCTCTCACTTATTAAAGATCTTGGAGCTGCTCAAGAGAGAGAGAGAGAGAGAGAG 1169
Qy 993 GCTGTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1052
Db 1170 GCTGTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1229
Qy 1053 GGTAAATATTTCCTCACTAGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1112
Db 1230 GGTAAATATTTCCTCACTAGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1289
Qy 1113 TGGAGTGAATTCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1172
Db 1290 TGGAGTGAATTCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1349
Qy 1173 GCTGTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232
Db 1350 GCTGTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1409
Qy 1233 TTTGATTCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1292
Db 1410 TTTGATTCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1469
Qy 1293 CTTTCATCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1352
Db 1470 CTTTCATCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1529
Qy 1353 TGTCTCTGTGTTGGGTGTTGTTCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1412
Db 1530 TGTCTCTGTGTTGGGTGTTGTTCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1589
Qy 1413 GTCC 1416

```

```

Db 1590 GTCC 1593
RESULT 11
AF131077 1726 bp mRNA linear ROD 03-JAN-2003
LOCUS Rattus norvegicus liver annexin-like protein (LAL) mRNA, complete cds.
DEFINITION AF131077.1 GI:7108712
ACCESSION AF131077
VERSION AF131077.1 GI:7108712
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 1726)
Della Fazio, M.A., Piobbico, D., Bartoli, D., Castell, M., Brancorsini, S., Viola Magni, M., and Servillo, G. 1999. A differentially expressed novel gene during proliferation in liver regeneration and in hepatoma cells. Genes Cells 7 (11), 1183-1190 (2002)
JOURNAL MEDLINE 22278398
PUBMED 12390252
AUTHORS Servillo, G., Della Fazio, M.A., Piobbico, D., Bartoli, D., Castell, M., Brancorsini, S., and Viola Magni, M.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1999) Institute of General Pathology, University of Perugia, Policlinico Monteluce, Perugia 06100, Italy
FEATURES
SOURCE 1. 1726
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1. 1726
/gene="LAL"
63..1481
/gene="LAL"
/note="Involved in liver regeneration; contains unique annexin domain at the N-terminus"
/codon_start=1
/product="liver annexin-like protein"
/protein_id="AAP6518.1"
/db_xref="GI:7108713"
/translaction="MRFLFLVAVVHLSELSGSKATYKSGVSGRTQRIKEIENYAE
DVAKAIINLAVYKQYRSEYERGLVDIVGPRISGSKLEKAIQIMYQMLQDGLN
VHLEOVYIPIHMERGESAVVVPRIHLALILGSGSIGTPEBGLTAEVLVASFVELO
RRASEARGKLVVNOPTDYKTVQYERGAVALKAVASLIRSVASFISVPHTG
HOGYODGVPRIPIACITIEDAEMSRASRSGDKIVHLKMGAKTYPTDSFNTVAEIT
GSKYPERVAVVSHLSDMDVQCALDDGGAFISWELSLVXDLGAPRKRTRLVMT
AESRGVGSQVYELHRAKANSKYSLVWEASDGFLLPTGLPTSDKRAIMKEVMSL
OPNLTIKVFNDAGTIDINFWIQGVDPASLRDLKYFFHSHGDTMTAMDPRKMNIV
AAAVMAVAVVAVVADMEMLPRS"
BASE COUNT 474 a 385 c 428 g 439 t
ORIGIN
Query Match 78.2%; Score 1107.2; DB 10; Length 1726;
Best Local Similarity 86.4%; Pred. No. 1,7e-305;
Matches 1223; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
Qy 1 ATGAATTCCTTATCTTCGATTTTGGGTGGCTTACCTTTATCCCTGTGTGGG 60
Db 63 ATGAGGTTCTTTCTTCTGTTGCTGTGCTGTGTTTACCTTTTCTCTGGGCTCGGA 122
Qy 61 AAAGCTATATCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 123 AAAGCTATATCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 182
Qy 121 GCCAGCTGTGAGAGATGTTGCTAAAGCATATCACTAGCTGTTATGTAAAGCCGAG 180
Db 183 GCCAATATGAAAGATGTTGCTAAAGCATATCACTAGCTGTTATGTAAAGATACGAG 242

```

QY 181 AACAGATCTATGAGCGATGGCACTTCTGTGATGATCTGTGGAGCCAGAGTGAAGG 240
 Db 243 AACCGCTGATAGAGCGTTTGGAGCTTCTGTGATGATCTGTGGAGCCAGAGTGAAGG 302
 QY 241 TCAGAGAACCTAGAAAAGGCAATCCAAATTAATGATCCAAAACCTGACGAATGGGCTG 300
 Db 303 TCTAAGAACCTAGAAAAGGCAATCCAAATTAATGATCCAAAACCTGACGAATGGGCTG 362
 QY 301 GAGAAAGTCACTGAGAGCGATGAGAAATCCCACTGGAGAGAGGAGAGAAATCAAGCT 360
 Db 363 GAAAACTCTCACTGGAGAGCGTCAAGAAATCTCACTGGAGAGAGGAGAGAAATCTGCA 422
 QY 361 GTGATGCTGAGCGCAAGAAATTAATGATGATCCCTGGGCTTGGAGAGAGCTTGG 420
 Db 423 GTGATGCTGAGCGCTCAAGAAATTAATGATGATGATGATGATGATGATGATGATG 482
 QY 421 ACTCTCTCAGAGAGCTTAACAGCAAGATCTGTGTGATGATCTTTCATGATGATGATG 480
 Db 483 ACTCTCTCAGAGAGCTTAACAGCAAGATCTGTGTGATGATCTTTCATGATGATGATG 542
 QY 481 AGAAGGCTCTCAGAGAGAGAGAGAGAGATGTTGTTTATTAACCACTTACATCACTAC 540
 Db 543 AGAAGGCTCTCAGAGAGAGAGAGAGATGTTGTTTATTAACCACTTACATCACTAC 602
 QY 541 TCAAGAGCGGTGCAATCCGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 Db 603 GGGAGAGCTGTGAG 662
 QY 601 GCAATCTCTATTCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGAT 660
 Db 663 GCAATCTCTATTCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGAT 722
 QY 661 GAATACCAAGATGCGGAG 720
 Db 723 GATATCAAGATGCGGAG 782
 QY 721 ATGATGTCAG 780
 Db 783 ATGATGTCAG 842
 QY 781 AAGAGCTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 Db 843 AAGAGCTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
 QY 841 CCAG 900
 Db 903 CCAG 962
 QY 901 ATGATGATGAG 960
 Db 963 CTGATGATGAG 1022
 QY 961 GAGCTGAG 1020
 Db 1023 GAGCTGAG 1082
 QY 1021 GTTGTGCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 Db 1083 GTTGTGCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1142
 QY 1081 ATGATGCTGAG 1140
 Db 1143 ATGATGCTGAG 1202
 QY 1141 GCGAG 1200
 Db 1203 GCGAG 1262
 QY 1201 CTGAG 1260
 Db 1263 TTTATGATGAG 1322

QY 1261 AGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 Db 1323 AGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1382
 QY 1321 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 Db 1383 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1442
 QY 1381 GTTGTGAG 1416
 Db 1443 GTTGTGAG 1478
 RESULT 12
 AX400829 1778 bp DNA linear PAT 06-JUN-2002
 LOCUS AX400829
 DEFINITION Sequence 505 from Patent WO0210453.
 ACCESSION AX400829
 VERSION AX400829.1 GI:21337009
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and
 Elashoff, M.R.
 TITLE Molecular toxicology modeling
 JOURNAL Patent: WO 0210453-A 505 07-FEB-2002;
 Gene Logic, Inc. (US)
 FEATURES
 source
 1..1778
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /note="EMBL/GenBank Accession No. AF097723"
 BASE COUNT 488 a 391 c 442 g 457 t
 ORIGIN
 Query Match 78.2%; Score 1107.2; DB 6; Length 1778;
 Best Local Similarity 86.4%; Pred. No. 1.7e-305;
 Matches 1223; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
 QY 1 ATGAATATCTTATCTTGGATTTTCCGATGATGATGATGATGATGATGATGATGATG 60
 Db 121 ATGAGGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 180
 QY 61 AAGCTATATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 Db 181 AAGCTATATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 121 GCCAGCTGAGAGATGTTGTTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 180
 Db 241 GCCAATATGAGATGTTGTTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 300
 QY 181 AACAGATCCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 Db 301 AACCGTGTATGAGAGCTTTGGAGCTTCTAGTTGATATCTGTGAGCCAGACTGAGTGG 360
 QY 241 TCCAGAGAACCTAGAAAAGGCAATCCAAATTAATGATCCAAAACCTGACGAATGGGCTG 300
 Db 361 TCTAAGAACCTAGAAAAGGCAATCCAAATTAATGATCCAAAACCTGACGAATGGGCTG 420
 QY 301 GAGAAAGTCACTGAG 360
 Db 421 GAAAGCTCACTGAG 480
 QY 361 GTGATGCTGAG 420
 Db 481 GTGATGCTGAG 540
 QY 421 ACTCTCTCAG 480

Db	541	ACTCTCCTGAAGGATATACAGAGAAAGTACTCGTGGGCGCTCTTTTGTTGAACCTTAA	600
QY	481	AGAAAGGCGCTCAGAAAGCAAGAGGAGATGTGTTATTAACCAACTTACATCACTAC	540
Db	601	AGAAAGGCGCATCAGAGGCAAGAGGAGATGTGTTATTAACCAAGCCTTACACTGACTAT	660
QY	541	TCMAAGACGGTGTGCATACCGAAACGAGAGGGGCGGTGGAAAGCTGCCAAGGTGGGGCTTGG	600
Db	661	GGGAAACATGTGTGCAGTACCCGGGAGCGGAGAGCTGTGGAAAGCTGCCAAGGTGGGGCGGTG	720
QY	601	GCATCTCATTCGATCCGTGGCGCTCTTCTTCATCTACAGTCCCTCACAGATATTGAG	660
Db	721	GCATCCCTCATCCGATTCAGTAGCTCTTTTTCATCTACAGTCCCTCACAGTATATGA	780
QY	661	GAATTCACGAGATGGCGTGCCCAAAATTCACACAGCCTGTATTAAGGTGAAGATGACAGA	720
Db	781	GGATATCAAGATGGGTGCGCCCAAGATTCACACAGCCTGTATTCACAAATGAAGATGACAGA	840
QY	721	ATGATGTCAAGAAATGGCTTCTTCAATGGGATCAAAATTTGTCACTTACGCTTAAAGTGGGGCA	780
Db	841	ATGATGTCTCAATATGGCTTCTGTGGGGACAAATTTGTCACTTACGCTTAAATGGGAGCA	900
QY	781	AAGACCTTACCCAGATACTGATTCCTTCAACACTGTAGAGAGATCACTCGGGAGCAAAATAT	840
Db	901	AAGACTTATCCAGATACAGATTCCTTCAACACTGTGACAGATCACTCGGGAGCAAGTAT	960
QY	841	CCAGAAACAGTTGTACTGCTGCAGTGGAGATCTGACAGCTGGGATGTTGGGCAAGGTC	900
Db	961	CCAGAGGAAGTTGTCTCGTGTGAGTGCATCTGACAGCTGGGAGTGTGGGCAAGGTC	1020
QY	901	ATGATGATATGGCGGTGGAGCCTTTATATCAATGGGAAGCACTCTCACTTATTAAGATCTT	960
Db	1021	CTGATGTATGGCGGTGGAGCCTTCAATCAATGAGGAAGCACTCTCACTTGTAAAGATCTT	1080
QY	961	GGGCGCGCTCCAAAGAGGACTGTGGGCGTGGCTCTGTAATGCGAAGAAACAAGGTGGA	1020
Db	1081	GGGCTGCGCTCCAAAGAGGACTGTGGGCGTGGCTCTGTAATGCGAAGAAACAAGGAGG	1140
QY	1021	GTGTGTGCTTCAAGTATTTATCAATGACACAAGGTAAATATTTCCAACTACAGTCTGGTG	1080
Db	1141	GTGTGTGCTTCCCAATATTTATGAGCTAATTAAGCAAAATTTTCCAAATGACGTTTGGTG	1200
QY	1081	ATGAGTCTTACGACGAGAACTTTCTTACCCACTGGGCTGCAATTCATCTGGCAGTGAAG	1140
Db	1201	ATGAGGCTGTGACTCAGGAACCTTCTTACCCACTGGGCTGCAATTCACCGGACGTGAAG	1260
QY	1141	GCCAGGGGCATATGAGAGGAGTATTAAGCCGTGACGCGCCCTCAATATCACTAGGTC	1200
Db	1261	GCCAGGGGCTATATGAGAGGAGTATTAAGCTTCTTGCACACCCCTCAATATCACTAAGGTC	1320
QY	1201	CTGAGCCATGAGAGGAGCAGACATCAACTTTTGATTCGAAGCTGAGTGCCTGAAGCC	1260
Db	1321	TTTATATGATGCAAGAGAACTGACATTTAACTTCTGTGATTCGAAGCTGAGTGCCTGAAGCC	1380
QY	1261	AGTCTACTTGAATGACTTATPAACAATTTCTTCTTCATCACTCCACGAGACACAATG	1320
Db	1381	AGTCTGCAAGATGACTTGTACAAATTTCTTTTCCATCAATCCCAATGAGACACAATG	1440
QY	1321	ACTGTCATGATCCAAAGCAAGATGTAATGTGTGTGCTGCTGTTGGGCTGTGTTCTTAT	1380
Db	1441	ACTGTCATGATCCAAAGCAAGATGTAATGTGTGTGCTGCTGTTGGGCTGTGCTTAT	1500
QY	1381	GTGTGTCAGACATGAGAAATGCTGCCAGGTC	1416
Db	1501	GTGTGTCAGACATGAGAAATGCTGCCAGGTC	1536

ACCESSION AF097723 GI:3851631
VERSION AF097723.1
KEYWORDS
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eulayotia, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
REFERENCE 1 (bases 1 to 1778)
AUTHORS Chen, Y. and Talmage, D.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1998) Institute of Human Nutrition, Columbia University, 701 West 168th Street Room 5-503, New York, NY 10032, USA

FEATURES
source
1.1778
Location/Qualifiers
1.1778
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="fisher"
/db_xref="taxon:10116"
/cell_line="F111"
/cell_type="fibroblast"
/dev_stage="embryo"
1.1778
/gene="H1s2-rp"
121.1539
/gene="H1s2-rp"
/codon_start=1
/product="hematopoietic lineage switch 2 related protein"
/protein_id="AAC72384.1"
/db_xref="gi:3851632"
/translation="MRPLFLFVAVVHLFSLGSKATIKSGVSRTPQRIKEINLY
DVAAIILNAVYGRQNSYERLGLVDVTPRISGSKNLEAIIMTONLQDQJEL
VHEOVRIPHWGRBESAVMVPVIRHKAIILGLGSGISCTPEGIITAEVLVASFEL
RRAESRGIKIVYNQPYTDYGTQVQYRGRGAVEAKVAVASLIRSVASFISYSHIT
HOGRODGPRIKTPACTITTEDAEMSRMRSGKIVIHLMGAKTVPDTSFTVAEIT
GSKRPVRYVYSGHLDSDVGCALDDCGGAFISWEALSLVQDYLRLPRRTILVLM
AEEQGVGASQYTELHKANIKRISLIMEADSGTPIPTGLQPTGSDPKARIMKEVNSL
QPLNITKVFNDABEGDINFWIQAGVPSLRDDLKYKFFPHSHDGTMTADPKQNN
AAAVAVAVVAVVAAMEMLPRS"
BASE COUNT 488 a 391 c 442 g 457 t
ORIGIN
Query Match 78.2%; Score 1107.2; DB 10; Length 1778;
Best Local Similarity 86.4%; Pred. No. 1.7e-305;
Matches 1223; Conservative 0; Mismatches 193; Indels 0; Gaps 0

1 ATGAATTCCTATTCCTGGCAATTTTGGGTGGTGTCACTTTTATCCCTGTGCTGGG 60
121 ATGAGTTCCTTTCTTCTCTGTTCTTGCTGTGTACCTTTTCTCTTGGGCTTGGA 180
61 AAAGCTATATCAAGATGGCATCTCTAAGAGGACTTTTGAAGAAATAAAGAGAATA 120
181 AAAGCTATATCAAGAGTGGTTCCTCAGGCAATTTCAAGAAATATAAGAGATA 240
121 GCGAGCTGTGAGATGTGCTTAAGCAATCATCAACTGACTGTTTATGTGTAAGCCG 180
241 GCCAAGTATGAAGAGTGTCTTAAGCAATATCAACTTGTGTTATGGAATAATCCAG 300
181 AACGATCTTATGAGCGATGGCACTTCCTGGTGTACTGTTGAGCCGACCTGAGTGGC 240
301 AACCGTGTATGAGCGTGTGGGACTTTAGTGTATGACTGTTGAGCCGACCTGAGTGGC 360
241 TCCAAGAACCTAGAAAAGCAATCCAAATTAATGTACCAAAACCTGACGCAATGGGCTG 300
361 TCTAAGAACCTAGAGAAAGCTATCCAAATCATGTATCCAAACCTGCACAAAGATGGGCTG 420
301 GAGAAAGTTCACCTGAGCCAGTGAAGATACCCCACTGGGAGAGGGGAGAGAATTCAGCT 360
421 GAAAAAGTTCACCTGAGCCAGTGAAGATACCTCACTGGGAGAGGGGAGAGAATTCAGCA 480
361 GTGATGCTGAGAGCCAAAGATTCATAGATAGCCATCTCTGGGTCTTGGCAGCAGCATTTGGG 420

```

Db 481 GTGATGCTGCTGCTCAATTCAAGTTGGCTATTTTAGCCCTTGGCCGAGCATTGGG 540
Qy 421 ACTCTCTCAGAGGCAATTAGACAGAGATTCTGGTGGTGAACCTCTTTCATGAACTGAG 480
Db 541 ACTCTCTCAGAGGCAATTAGACAGAGATTCTGGTGGTGAACCTCTTTCATGAACTGAG 600
Qy 481 AGAAGGCTCTCAGAGGCAATTAGAGAGATTGTTTATTAACAACCTTACATCACTAC 540
Db 601 AGAAGGCTCTCAGAGGCAATTAGAGAGATTGTTTATTAACAACCTTACATCACTAC 660
Qy 541 TCAGAGAGGCTGCAATTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db 661 GGGAGAACTGTGAGTACCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Qy 601 GCATCTCTCATGAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 721 GCATCTCTCATGAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Qy 661 GAAATACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Db 781 GATATATCAAGATGGTGGTCCCAAGATTCCAAGAGCTGTATCAATATGAAGATGAGAA 840
Qy 721 ATGATGTCAGAGATGGTCTTCTCATGAGATCAAAATTTGCTATCAGCTAAAGATGGG 780
Db 841 ATGATGTCAGAGATGGTCTTCTCATGAGATCAAAATTTGCTATCAGCTAAAGATGGG 900
Qy 781 AAGACCTACCAAGATGATGATCTTCAACAGCTGTAGCAGAGATCACTGGGAGCAATAT 840
Db 901 AAGACCTACCAAGATGATGATCTTCAACAGCTGTAGCAGAGATCACTGGGAGCAATAT 960
Qy 841 CCAGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 961 CCAGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Qy 901 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 1021 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Qy 961 GGGCTGCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 1081 GGGCTGCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Qy 1021 GTTGTGCTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1141 GTTGTGCTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Qy 1081 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1201 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Qy 1141 GCGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Db 1261 GCGAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Qy 1201 CTGAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1321 TTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
Qy 1261 AGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1381 AGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Qy 1321 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
Db 1441 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Qy 1381 GTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Db 1501 GTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560

```

RESULT 14
AF009513

```

LOCUS AF009513 1716 bp mRNA linear ROD 01-JUN-1998
DEFINITION Mus musculus hematopoietic lineage switch 2 (HLS2) mRNA, complete
cde.
ACCESSION AF009513
VERSION AF009513.1 GI:3169728
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITL
JOURNAL
REFERENCE
AUTHORS
TITL
JOURNAL
FEATURES
source
1. 1716
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1. 1716
/gene="HLS2"
97. 1398
/gene="HLS2"
/product="hematopoietic lineage switch 2"
/protein_id="AA17945.1"
/db_xref="GI:3169728"
/translaction="MRSLEFLPLVHLALSGKAVKNGVORTFREIKEIANYEDV
AKAIINLAVYKXONRYERGLVDTPVPRSGSKLEKAIQIOMQNLQODLENVH
LEQVRI.PHMERGESAVMLEPRHMAIIGLSSIGTPGGITAEVLVAASFELGR
ASEBARKIIVNOPRYGKTYOYRQVGAKEAIVKAVASLIQSVASFISPHRTIG
KXODGPKIPTVACITVEDAEKMSRMSRKNKIVIHLEMAKTYPDSDNTVAETIS
MPEDVGQALDDGGAFTSWEALSLVKOLGPRKPTLRLVMTAEQSGTGSQYE
LHRANISKSLVNEADSGFLPLGLOFTSDKARAIKEMVNLLOPLNTVYFSNGBG
TDINFWIQGVGASLRDLDLYKFFPHSHGDJMTMSIOSR"
BASE COUNT 476 a 370 c 414 g 456 t
ORIGIN
Query Match 72.8%; Score 1030.8; DB 10; Length 1716;
Best Local Similarity 85.6%; Pred. No. 1.3e-283;
Matches 1185; Conservative 0; Mismatches 162; Indels 37; Gaps 2;
Qy 33 TGTTCACCTTTTATCCCTGTGCTCTGAGAAAGCTATATGCAAAAGTCATCTTAAG 92
Db 123 TGTTCACCTTTTATCCCTGTGCTCTGAGAAAGCTATATGCAAAAGTCATCTTAAG 162
Qy 93 GACTTTTGAAGAATATAAGAAATAGCCAGCTGTGAGATGTTGCTTAAAGCAATCAT 152
Db 183 AACATTTCCAGAAATATAAGAAATAGCCAGCTGTGAGATGTTGCTTAAAGCAATCAT 242
Qy 153 CAACCTAGCTGTTTATGTTAAGCCAGAAACAGATCCATATGACGATGACCTTCTGT 212
Db 243 CAACCTAGCTGTTTATGTTAAGCCAGAAACCGGCTCATATGACGATGACCTTCTGT 302
Qy 213 TGATACCTGTTTGAAGCCAGAGCTGAGTGCCTCCAGAACTTGAAGAAAGCCATCCAAATAT 272
Db 303 TGATACCTGTTTGAAGCCAGAGCTGAGTGCCTTGAAGAACTTGAAGAAAGCCATCCAAATAT 362
Qy 273 GTACCAAAACCTGACAGAAAGATGGGCTGAGAAAGTTCACTGAGCAGTGAATAC 332
Db 363 GTACCAAAACCTGACAGAAAGATGGGCTGAGAAATGTTCACTGAGCAGTGAATAC 422
Qy 333 CCACCTGGAGAGGGGAGAAAGATCAGCTGTGATGCTGAGCCAAAGATTCAATAGATAG 392
Db 423 CCACCTGGAGAGGGGAGAAAGATCAGTGAATGCTGAGCCCTGAATTCACAAAGATGAG 482

```

```

OY 393 CATCTGGGCTTGGCAGACGATGGAGCTCTCCAGAAAGCATTTACAGAGAACTTCT 452
Db 483 TATCTAGTGTCTTGGCAGACGATGGAGCTCTCCAGAAAGCATTTACAGAGAACTTCT 542
OY 453 GGTGGTGAAGCTCTTTCGATGAACTGACAGAAAGGGCTCAGAAAGCAGAGAAATGTT 512
Db 543 GGTGGTGAAGCTCTTTCGATGAACTTCAAAAGAAAGCATTTAGAAAGCAGAGAAATGTT 602
OY 513 TGTTTTAAACCACTTTATCATCACTACTCAAGAGCGGTGCAATTCGAAAGCGAGGGGG 572
Db 603 TGTTTTAAACCACTTTATCACTGCTATGAGAAAGCTGTGACAGTACCGGGTGGAGGAGC 662
OY 573 GGTGAAGCTCTCAAGGTGGGGGCTTGGGCACTCTCATTCGATCCGTGGCTCTCTTC 632
Db 663 TGTGAAGCTCTCAAGGTGGAGGTGGAGCTGTGGCATCTCTCATTCAGTAAAGTCTCTTTC 722
OY 633 CATCTACAGTCTCTCACACAGGTATTCAAGAAATCCAGAGTGGCGTCCCAAAATTCAC 692
Db 723 CATCTACAGTCTCTCACACAGGTATTCAAGAAATCCAGAGTGGTGGCCCAAGATTCAC 782
OY 693 AGCCTGTATTACGGTGAAGATGACAAATATGTCAGAAATGCTTCTCATGGGATCAA 752
Db 783 AGCCTGTATTACAGTGAAGATGACAAATATGTCAGAAATGCTTCTCTCGTGGAAACAA 842
OY 753 AATTGTCAATCAGCTAAAGATGGGGGCAAAAGCTACCCAGATCTGATTCCTTCAACAC 812
Db 843 AATTGTCAATCAGCTGAAGATGGAGCAAAAGCTATTCAGATCTGATTCCTTCAATAC 902
OY 813 TGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTTGTACTGATGACATCT 872
Db 903 TGTAGCAGAGATCACTGGGAGCAATGATATCCAGAA----- 935
OY 873 GACAGCTGGATGTTGGGAGGGGTCAGATGATGATGGCGTGGAGGCTTATATCATG 932
Db 936 -----GATGTTGGGAGGGGTCAGATGATGATGGCGTGGAGGCTTATATCATG 986
OY 933 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGGCTCAAAAGAGACTTCGGGCTGT 992
Db 987 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGGCTCAAAAGAGACTTCGGGCTGT 1046
OY 993 GCTCTGACCTGACAGAAAGCAAGGTGAGTGGTCCCTTCAGATTTATCAGATTACAA 1052
Db 1047 GCTCTGACCTGACAGAAAGCAAGGTGAGTGGTCCCTTCAGATTTATGACATCAATA 1106
OY 1053 GGTAAATTTTCCAACTACAGTCTGTGATGAGAGTCTGACGAGAACTCTTACCCAC 1112
Db 1107 GGTAAATTTTCCAACTACAGTCTGTGATGAGAGTCTGACGAGAACTCTTACCCAC 1166
OY 1113 TGGGCTGCAATTTCACTGGCAGTGAAGAAAGCCAGGGCCATCATGAGAGGTTATGAGCT 1172
Db 1167 TGGGCTGCAATTTCACTGGCAGTGAAGAAAGCCAGGGCCATCATGAGAGGTTATGAGCT 1226
OY 1173 GCTGACGCTCTCAATATCACTCAGTCTGAGCCATGAGAGAGGAGACATCAACTT 1232
Db 1227 TCTGGAACCTCTCAATATCACTCAGTCTGAGCCATGAGAGAGGAGTATGATTAATT 1286
OY 1233 TTTGATCCAGCTGAGAGTGGCTGAGAGCAAGTCTATTGATGACTTATTAAGTATTTCT 1292
Db 1287 CTGATTCAGCTGAGAGTGGCTGAGAGCAAGTCTGATGATGACTTATTAAGTATTTCT 1346
OY 1293 CTTTCATCACTCTCCAGAGAGACACCATGATCTGTGATCCAAAGCAGATGATGTTGC 1352
Db 1347 TTTTCATCACTCTCCAGAGAGACACCATGATCTGTGATCCAAAGCAGATGATGTTGC 1405
OY 1353 TGTGCTGTTTGGGCTGTTGTTTCTTATGTTGTCAGACATGAGAAAGTCTGCTAG 1412
Db 1406 TGTGCTGTTTGGGCTGTTGTTGCTTATGTTGTCAGACATGAGTAAAGTCTGCCAG 1465
OY 1413 GTCC 1416
Db 1466 GTCC 1469

```

```

RESULT 15
BD059610 609 bp DNA linear PAT 27-AUG-2002
DEFINITION Secreted expressed sequence tags (ESTs).
LOCUS BD059610
ACCESSION BD059610.1 GI:22605216
VERSION JP 2001519666-A/1465
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 609)
REFERENCE
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Werberg,D.,
Treacy,M., Spaulding,V. and Agostino,M.J.
Secreted expressed sequence tags (ESTs)
Patent: JP 2001519666-A/1465 23-OCT-2001;
GENETICS INSTITUTE INC
PN JP 2001519666-A/1465
PD 23-OCT-2001
PR 10-APR-1998 JP 1998543068
PI 10-APR-1997 US 08/835913
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:
Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source Location/Qualifiers
1..609
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
BASE COUNT 190 a 127 c 148 g 144 t
ORIGIN
Query Match 36.8%; Score 520.6; DB 6; Length 609;
Best Local Similarity 99.1%; Pred. No. 1,be-137;
Matches 534; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
OY 1 ATGAATTCCTTATCTTGGATTTTTCGATGTTGTCACCTTTATCCGTGCTGGG 60
Db 71 ATGAATTCCTTATCTTGGATTTTTCGATGTTGTCACCTTTATCCGTGCTGGG 130
OY 61 AAAGCTATATGCAAGATGCGATCTTAAAGAGACTTTTGAAGAAATAAAGAAATA 120
Db 131 AAAGCTATATGCAAGATGCGATCTTAAAGAGACTTTTGAAGAAATAAAGAAATA 190
OY 121 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 180
Db 191 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 250
OY 181 AACAGATCTATGAGCGATTTGGCACTTCTGTTGATCTGTGAGCCAGCTGAGTGGC 240
Db 251 AACAGATCTATGAGCGATTTGGCACTTCTGTTGATCTGTGAGCCAGCTGAGTGGC 310
OY 241 TCCAGAACTTGAAGAAAGCCATTCAAATTTATGTAACAAAACCTGCGCAAGATGGCTG 300
Db 311 TCCAGAACTTGAAGAAAGCCATTCAAATTTATGTAACAAAACCTGCGCAAGATGGCTG 369
OY 301 GAGAAAGTTCACTGGAGCGCAGTGAATAACCCCACTGGGAGAGGGGAGAAATCAGCT 360
Db 370 GAGAAAGTTCACTGGAGCGCAGTGAATAACCCCACTGGGAGAGGGGAGAAATCAGCT 429
OY 361 GTGATGCTGAGCGCAAGATTTCAATAGATAGCCATCTCGGTCTTGGCAGAGATGGG 420
Db 430 GTGATGCTGAGCGCAAGATTTCAATAGATAGCCATCTCGGTCTTGGCAGAGATGGG 489
OY 421 ACTCTTCAAGAGGCAATTAAGAGAAAGTTCTGTTGATCTTTTGAATGAATGAG 480
Db 490 ACTCTTCAAGAGGCAATTAAGAGAAAGTTCTGTTGATCTTTTGAATGAATGAG 549

```

Oy 481 AGAAGGGCCCTAGAAGCAAGAGGGAGAATTGTTTATTAACCAACTTAACTCACTA 539
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 550 AGAAGGGCCTCAGAGCAAGAGGAAGATTGTTTATTAACCAACTTAACTCGA 608

Search completed: December 22, 2003, 20:22:18
Job time : 5117.35 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 12:51:18 / Search time 341.284 Seconds
(without alignments)
11200.072 Million cell updates/sec

Title: US-09-745-763-35_COPY_99_1514

Perfect score: 1416
Sequence: 1 ATGAATTCCTTATCTTCG.....AGAAATGCTGCTAGTCC 1416

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1416	100.0	1851	19	AAV82779
2	1416	100.0	1851	24	ABO92016
3	1412.8	99.8	1767	21	AAA04093
4	1412.8	99.8	1778	18	AAV02296
5	1412.8	99.8	1863	21	AAZ98034
6	1412.8	99.8	1863	22	AAD11647
7	1412.8	99.8	1863	24	ABK69743
8	1412.8	99.8	1863	25	ACC50817

9	1412.8	99.8	1863	25	AB271453
10	1412.8	99.8	1884	21	AAZ58313
11	1412.8	99.8	1923	21	AAZ98139
12	1412.8	99.8	2077	25	ACC50603
13	1412.8	99.8	2077	25	AB271351
14	1408	99.4	1860	22	AAK94491
15	1355.6	95.7	1895	22	AAH9703
16	1296.8	91.6	1784	23	AAH35592
17	1107.2	78.2	1778	24	ABK62598
18	862.8	59.5	1895	21	AAA44359
19	772.4	54.5	1134	21	AAZ98101
20	772.4	54.5	1134	22	AAD11714
21	772.4	54.5	1134	24	ABK69810
22	772.4	54.5	1134	25	ACC50818
23	772.4	54.5	1134	25	AB271454
24	520.6	36.8	609	20	AAV87487
25	470.2	33.2	642	22	AAK92181
26	470.2	33.2	642	22	AAK93447
27	284.8	20.1	314	20	AAV86369
28	276.6	19.5	462	18	AAV97398
29	276.6	19.5	462	18	AAV88060
30	276.6	19.5	462	18	AAV02139
31	242.4	17.1	441	22	AAK56831
32	212.6	15.0	317	20	AAK40587
33	114.2	8.1	492	22	AAK92932
34	90.8	5.4	394	14	AAO61260
35	84	5.9	217	20	AAK40583
36	61.4	4.3	424	23	AAK73591
37	60.2	4.3	65	24	ABN53183
38	60	4.2	60	24	ABN36212
39	42	3.0	784	22	AAH05496
40	42	3.0	2200	22	AAH78055
41	42	3.0	2200	22	AAH78055
42	42	3.0	2408	22	AAH99550
43	41	2.8	65	24	ABN28454
44	39.4	2.8	975	21	AACT7173
45	39.4	2.8	1556	24	AAJ33726

ALIGNMENTS

AAV82779	standard; cDNA; 1851 BP.
AAV82779;	
25-FEB-1999	(first entry)
DE	Clone bu45_2 isolated from human adult placenta cDNA library.
XX	Secreted protein; nutritional activity; immune stimulating; vaccine;
KW	suppressing activity; haematopoiesis regulating activity;
KW	tissue growth activity; activity; inhibin activity; chemotaxis;
KW	chemokinetic activity; haemostasis; thrombolytic activity; receptor;
KW	ligand; anti-inflammatory; cadherin; tumour invasion suppressor;
XX	tumour inhibition; gene therapy; ds.
OS	Homo sapiens.
XX	
FN	MO9842739-A2.
XX	
PD	01-OCT-1998.
XX	
PF	20-MAR-1998; 98MO-US05653.
XX	
PR	19-MAR-1998; 98US-0044466.
XX	
PR	21-MAR-1997; 97US-0822167.
XX	
PA	(GENY) GENETICS INST INC.
XX	
PI	Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding V, Treacy M;
 XX WPI: 1998-609890/51.
 DR P-PSDB; AAM85456.
 XX New polynucleotides encoding secreted human proteins - derived from
 PT human foetal brain, adult brain, foetal kidney, placenta or adult
 PT pineal gland cDNA libraries.
 XX
 PS Claim 14; Page 69-70; 113pp; English.
 CC The present sequence encodes a secreted protein. The polynucleotide and
 CC secreted protein are predicted to have biological activities which would
 CC make them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is given.
 CC Suggested activities include nutritional activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity (no data is
 CC given in the specification to support these activities). The
 CC polynucleotide is also stated to be useful for gene therapy.
 XX
 SQ Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;
 Query Match 100.0%; Score 1416; DB 19; Length 1851;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAATTCCTATCTTGCATTTTCGGTGGTTCACCTTTATCCCTGCTGGG 60
 99 ATGAATTCCTATCTTGCATTTTCGGTGGTTCACCTTTATCCCTGCTGGG 158
 61 AAAGCTATATGCAAGATGGCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATA 120
 159 AAAGCTATATGCAAGATGGCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATA 218
 121 GCCAGCTGTGAGATGTTGCTTAAAGCATCATCACTAGCTTTATGTTAAAGCCAG 180
 219 GCCAGCTGTGAGATGTTGCTTAAAGCATCATCACTAGCTTTATGTTAAAGCCAG 278
 181 AACAGATCTTATGAGCGATTTGGCACTTGGTTGATCTGTGAGCCAGCTGAGTGGC 240
 275 AACAGATCTTATGAGCGATTTGGCACTTGGTTGATCTGTGAGCCAGCTGAGTGGC 338
 241 TCCAAGAACTTGAAGAAAGCCATCCAAATTAATGTAACCAAACTTGCAGCAAGATGGGCTG 300
 339 TCCAAGAACTTGAAGAAAGCCATCCAAATTAATGTAACCAAACTTGCAGCAAGATGGGCTG 398
 301 GAGAAAGTTCACTTGGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAAGAACTCACT 360
 395 GAGAAAGTTCACTTGGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAAGAACTCACT 458
 361 GTGATGCTGAGCCAGAAATTCATAAGATAGCCATCTGGGCTTGGCAGCAGCTTGGG 420
 455 GTGATGCTGAGCCAGAAATTCATAAGATAGCCATCTGGGCTTGGCAGCAGCTTGGG 518
 421 ACTCTTCAGAGAGGATTAACAGCAGAGTTCTGTGTGATCACTTTTGATTAACCTGAG 480
 519 ACTCTTCAGAGAGGATTAACAGCAGAGTTCTGTGTGATCACTTTTGATTAACCTGAG 578
 481 AGAAGGGCTCAGAGAGGAGAAATGTTGTTTATTAACCAACTTATCACTCACTAC 540
 579 AGAAGGGCTCAGAGAGGAGAAATGTTGTTTATTAACCAACTTATCACTCACTAC 638
 541 TCAAGAGCGTGCATTAACGAGAGCGAGGGCGGTGAGAGCTGCCAAGGTGGGGCTTGG 600
 639 TCAAGAGCGTGCATTAACGAGAGCGAGGGCGGTGAGAGCTGCCAAGGTGGGGCTTGG 698
 601 GCATCTCTCATTCGATCCGCTGCTCTTCTTCATCTCAAGTCTCAGACAGATATTCAG 660
 699 GCATCTCTCATTCGATCCGCTGCTCTTCTTCATCTCAAGTCTCAGACAGATATTCAG 758

661 GAATACCAGATGGCGGCGCCCAAAATTCACACAGCCTGTATTAACGGTGAAGATGCAGAA 720
 759 GAATACCAGATGGCGGCGCCCAAAATTCACACAGCCTGTATTAACGGTGAAGATGCAGAA 818
 721 ATGATGTCAGAAATGCGTCTTCATGAGGATCAAAATTTGATTCATGCTAAAGATGGGGCA 780
 819 ATGATGTCAGAAATGCGTCTTCATGAGGATCAAAATTTGATTCATGCTAAAGATGGGGCA 878
 781 AAGACCTAACCCAGATTAATGATTTCTTCAACACTGTATACAGATCACTGGAGCAATAT 840
 879 AAGACCTAACCCAGATTAATGATTTCTTCAACACTGTATACAGATCACTGGAGCAATAT 938
 841 CCAGAACAGGTTGATGCTGATGAGCATCTGAGCAGCTGGAGTGTGGGAGGGTGGC 900
 939 CCAGAACAGGTTGATGCTGATGAGCATCTGAGCAGCTGGAGTGTGGGAGGGTGGC 998
 901 ATGAGATATGCGCGGTGAGGCTTTATATCATGAGGAGCACTTCACTTATTAAGATCTT 960
 999 ATGAGATATGCGCGGTGAGGCTTTATATCATGAGGAGCACTTCACTTATTAAGATCTT 1058
 961 GGGTGGCTCCAAAGAGACTCTGCGGCTGCTCTGAGCTGAGCAGAAAGCAAGGTGGA 1020
 1059 GGGTGGCTCCAAAGAGACTCTGCGGCTGCTCTGAGCTGAGCAGAAAGCAAGGTGGA 1118
 1021 GTTGGTGCCTCCAGTATTAATGATTAACAGAGTAAATATTTCACTCACTCACTGCTG 1080
 1119 GTTGGTGCCTCCAGTATTAATGATTAACAGAGTAAATATTTCACTCACTCACTGCTG 1178
 1081 ATGAGATCTGACGAGCAACCTTCTTACCACTGGGCTGCAATTCATGSCAGTGAAG 1140
 1179 ATGAGATCTGACGAGCAACCTTCTTACCACTGGGCTGCAATTCATGSCAGTGAAG 1238
 1141 GCCAGGGCCATCATGAGAGGTTATGAGCTGCTGAGCCCTCAATATCATCTCACTG 1200
 1239 GCCAGGGCCATCATGAGAGGTTATGAGCTGCTGAGCCCTCAATATCATCTCACTG 1298
 1201 CTGAGCCATGAGAGAGGAGACATCACTTTTGAATCCAGGTGAGTGGCTGAGGCC 1260
 1299 CTGAGCCATGAGAGAGGAGACATCACTTTTGAATCCAGGTGAGTGGCTGAGGCC 1358
 1261 AGTCTACTGATGATCACTTATCAAGATATTTCTTCCATCACTCCACGAGACACCATG 1320
 1359 AGTCTACTGATGATCACTTATCAAGATATTTCTTCCATCACTCCACGAGACACCATG 1418
 1321 ACTGTCTGATGATCCAAAGCAGATGATGTTGCTGCTGTTGGGCTGTTGTTCTTAT 1380
 1419 ACTGTCTGATGATCCAAAGCAGATGATGTTGCTGCTGTTGGGCTGTTGTTCTTAT 1478
 1381 GTTGTGAGCAATGGAAGAAATGCTGCTAGGTC 1416
 1479 GTTGTGAGCAATGGAAGAAATGCTGCTAGGTC 1514

RESULT 2
 AB092016 standard; cDNA; 1851 BP.
 ID AB092016;
 XX
 AC AB092016;
 XX
 DT 04-OCT-2002 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 13.
 XX
 XX Human; cytostatic; antirheumatic; antiarthritic; vulnery; analgesic;
 KW antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
 KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
 KW antidiuretic; fungicide; antidiabetic; antiaesthetic; antiallergic;
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KW stem cell; growth factor; nervous system disease; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;

KM multiple sclerosis, rheumatoid arthritis, gene therapy; gene; ss.
 XX Homo sapiens.
 OS
 XX US2002065394-A1.
 PN
 XX 30-MAY-2002.
 PD
 XX 22-DEC-2000; 2000US-0745763.
 PF
 XX 18-MAR-1998; 98US-0040963.
 PR
 XX (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (SPAU/) SPAULDING V.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 XX
 XX WPI; 2002-582343/62.
 DR P-PSDB; ABP61800.
 DR
 XX
 PT Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis
 PT
 PS Claim 50; Page 113-114; 284pp; English.
 XX
 XX The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (CDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (I) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, anti-inflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gene protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polynucleotide of the invention.
 CC
 XX Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;
 SQ
 Query Match 100.0%; Score 1416; DB 24; Length 1851;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAATTCCTTATCTTGCAATTTTCGGTGTGTTACCTTTATCCCTGTGCTGGG 60

Db
 99 ATGAATTCCTTATCTTGCAATTTTCGGTGTGTTACCTTTATCCCTGTGCTGGG 158
 QY
 61 AAAGCTATATGCAAGAAATGGCATCTCTAAGAGACCTTTGAAGAAATGAAGAAATA 120
 Db
 159 AAAGCTATATGCAAGAAATGGCATCTCTAAGAGACCTTTGAAGAAATGAAGAAATA 218
 QY
 121 GCCAGCTGTGGAAGATGTGTCTAAAGCAATCATCACTAGCTGTTATGGTAAAGCCAG 180
 Db
 219 GCCAGCTGTGGAAGATGTGTCTAAAGCAATCATCACTAGCTGTTATGGTAAAGCCAG 278
 QY
 181 AACAGATCTATAGAGCATTTGGCACTTCTGTGTGAATCTGTGGACCCAGCTGAGTGGC 240
 Db
 279 AACAGATCTATAGAGCATTTGGCACTTCTGTGTGAATCTGTGGACCCAGCTGAGTGGC 338
 QY
 241 TCCAGAACTTAAGAAAAAGCCATTCGAATTTATGACCAAAAACCTGACGCAAGATGGGCTG 300
 Db
 339 TCCAGAACTTAAGAAAAAGCCATTCGAATTTATGACCAAAAACCTGACGCAAGATGGGCTG 398
 QY
 301 GAGAAAGTTCACTGAGAGCCAGTGAATAATCCCACTGGGAGGAGGAGAAATCAGCT 360
 Db
 399 GAGAAAGTTCACTGAGAGCCAGTGAATAATCCCACTGGGAGGAGGAGAAATCAGCT 458
 QY
 361 GTGATGCTGAGAGCCAGAAATTCATATAGATAGCATCTGGGCTTTGGCAGACGATTTGGG 420
 Db
 459 GTGATGCTGAGAGCCAGAAATTCATATAGATAGCATCTGGGCTTTGGCAGACGATTTGGG 518
 QY
 421 ACTCTTCAGAGGACATTAACGACGAAGTTCTGTGTGATCCTTTTCATGAACTGACG 480
 Db
 519 ACTCTTCAGAGGACATTAACGACGAAGTTCTGTGTGATCCTTTTCATGAACTGACG 578
 QY
 481 AGAAGGCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
 Db
 579 AGAAGGCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 638
 QY
 541 TCAAGACGATGCAATACCAAGCAGAGGGGCGGTGAGACCTGCAAGATGGGGCTTTG 600
 Db
 639 TCAAGACGATGCAATACCAAGCAGAGGGGCGGTGAGACCTGCAAGATGGGGCTTTG 698
 QY
 601 GCATCTCTCATTTGATCGTGGCTCTCTTCATCTACAGTCTTCAACAGATGATTCAG 660
 Db
 699 GCATCTCTCATTTGATCGTGGCTCTCTTCATCTACAGTCTTCAACAGATGATTCAG 758
 QY
 661 GAATACAGGATGGCGGCCAAATTCGAAGCCGTGTTACGGTGAAGATGACGAA 720
 Db
 759 GAATACAGGATGGCGGCCAAATTCGAAGCCGTGTTACGGTGAAGATGACGAA 818
 QY
 721 ATGATGTCAGAAATGAGCTTCTCATGAGATCAAAATTTGCTAATTCAGTAAAGATGGGGCA 780
 Db
 819 ATGATGTCAGAAATGAGCTTCTCATGAGATCAAAATTTGCTAATTCAGTAAAGATGGGGCA 878
 QY
 781 AAGACTTACCAAGATGATGATTTCTTCAACACTGTAGACAGATCACTGGAGCAAAATAT 840
 Db
 879 AAGACTTACCAAGATGATGATTTCTTCAACACTGTAGACAGATCACTGGAGCAAAATAT 938
 QY
 841 CCAGAAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 Db
 939 CCAGAAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 998
 QY
 901 ATGATGATGAGCGGTGAGGCTTTATATCATGAGGAACACTCTCACTAATTAAGATCTT 960
 Db
 999 ATGATGATGAGCGGTGAGGCTTTATATCATGAGGAACACTCTCACTAATTAAGATCTT 1058
 QY
 961 GGGCTGCTCCAAAGAGAGACTCTGGGCTGGTCTGTGACTGACGAGAAAGCAAGGTGGA 1020
 Db
 1059 GGGCTGCTCCAAAGAGAGACTCTGGGCTGGTCTGTGACTGACGAGAAAGCAAGGTGGA 1118
 QY
 1021 GTTGTGCTCTTCAAGATTAATGATTAACAGGTAATTAATTTCCACTCAAGCTGGTG 1080
 Db
 1119 GTTGTGCTCTTCAAGATTAATGATTAACAGGTAATTAATTTCCACTCAAGCTGGTG 1178
 QY
 1081 ATGAGATCTGACGAGGAACTTTCTTACCACTGGGCTGCAATTCATGCGACGTGAAG 1140

Db 1179 ATGAGTCTGACGACGAACTTCTTACCCTAGGCTGCAATTCATGCGAGTGAAG 1238
 Qy 1141 GCCAGGGCCATCATGAGAGGTTATGAGCCTGCTGACGCCCTTCATATGACTCAGGTC 1200
 Db 1239 GCCAGGGCCATCATGAGAGGTTATGAGCCTGCTGACGCCCTTCATATGACTCAGGTC 1238
 Qy 1201 CTGAGCCATGAGAGGAGACAGACATCACTTTTGGATTCAGAGCTGAGTCCCTGAGCC 1260
 Db 1299 CTGAGCCATGAGAGGAGACAGACATCACTTTTGGATTCAGAGCTGAGTCCCTGAGCC 1358
 Qy 1261 AGTCTACTGATGATTAATACAGATTTTCTTCTTCATCACTCCACGAGACACCATG 1320
 Db 1359 AGTCTACTGATGATTAATACAGATTTTCTTCTTCATCACTCCACGAGACACCATG 1418
 Qy 1321 ACTGTCATGATTCACAAAGCAGATGATGCTGCTGCTGTTGGGCTGTTCTTAT 1380
 Db 1415 ACTGTCATGATTCACAAAGCAGATGATGCTGCTGCTGTTGGGCTGTTCTTAT 1478
 Qy 1381 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1416
 Db 1475 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1514

RESULT 3

AAAA0493 standard; cDNA; 1767 BP.

AC AAA0493;

DT 16-NOV-2000 (first entry)

XX Human fetal kidney cDNA fragment AM282_1f.

KW Secreted protein; cytosolic; immunostimulatory; antimicrobial;
 KW antiviral; immunosuppressive; antiinflammatory; vulnery; cytokine;
 KW cell proliferation; differentiation; regulation; treatment; tumor;
 KW autoimmune disease; inflammatory disorder; wound; microbial infection;
 KW viral disease; graft versus host reaction suppression; ss.

XX Homo sapiens.

XX WO200037630-A1.

XX 29-JUN-2000.

XX 22-DEC-1999; 99WO-US31005.

XX 23-DEC-1998; 98US-0220876.

XX (GENY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Metberg D, Treacy M, Bowman MR;

XX MPI; 2000-442661/38.

XX P-PSDB; AAB10229.

XX Secreted human proteins AS296-1f and AS34-1f, useful for treating

XX infections and viral diseases -

XX Disclosure; Page 198; 29pp; English.

CC This invention describes novel secreted human proteins (I) which have
 CC cytosolic, immunostimulatory, antimicrobial, antiviral,
 CC immunosuppressive, antiinflammatory and vulnery activity and which act
 CC as cytokine, cell proliferation or differentiation regulators. (I)
 CC is useful for treating tumors, autoimmune diseases, inflammatory
 CC disorders, wounds, microbial infections and viral diseases. (I) is also
 CC useful for suppressing graft versus host reaction. AAA0490-A40580
 CC represent cDNA fragments that encode the secreted proteins
 CC AAB10226-B10288 described in the method of the invention.

SQ Sequence 1767 BP; 512 A; 384 C; 418 G; 453 T; 0 other;

Query Match 99.8%; Score 1412.8; DB 21; Length 1767;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAATTCCTTATCTTGCAATTTTGGTGGTTCACCTTTATCTCTGCTGGG 60
 Db 6 ATGAATTCCTTATCTTGCAATTTTGGTGGTTCACCTTTATCTCTGCTGGG 65
 Qy 61 AAAGCTTATGCAAGATGGCATCTTAAAGACCTTTGAAAGAAATGAAGAAATA 120
 Db 66 AAAGCTTATGCAAGATGGCATCTTAAAGACCTTTGAAAGAAATGAAGAAATA 125
 Qy 121 GCCAGCTGTGAGATGTGTGCTTAAAGCAATCACTAGCTGTTATGTTAAAGCCAG 180
 Db 126 GCCAGCTGTGAGATGTGTGCTTAAAGCAATCACTAGCTGTTATGTTAAAGCCAG 185
 Qy 181 AACAGATCTATGAGCCATTTGGCACTTCTGTTGATACTGTGAGCCAGACTAGTGC 240
 Db 186 AACAGATCTATGAGCCATTTGGCACTTCTGTTGATACTGTGAGCCAGACTAGTGC 245
 Qy 241 TCCAAAGACTTAAAGAAAGCCATTCAAATTTATGTCACAAACCTGACGAAATGGGCTG 300
 Db 246 TCCAAAGACTTAAAGAAAGCCATTCAAATTTATGTCACAAACCTGACGAAATGGGCTG 305
 Qy 301 GAGAAAGTTCACCTGAGCCAGTGAATATCCCACTGGGAGAGAGAGAAATACAGCT 360
 Db 306 GAGAAAGTTCACCTGAGCCAGTGAATATCCCACTGGGAGAGAGAGAAATACAGCT 365
 Qy 361 GTGATGTGAGCCAGAAATTCATAGATGCCATCTGGGCTTGGCAGACGATTTGG 420
 Db 366 GTGATGTGAGCCAGAAATTCATAGATGCCATCTGGGCTTGGCAGACGATTTGG 425
 Qy 421 ACTCTTCAAGAGCATTTACGACGAAGTTCTGTGTGACCTCTTTGATGAACTGAG 480
 Db 426 ACTCTTCAAGAGCATTTACGACGAAGTTCTGTGTGACCTCTTTGATGAACTGAG 485
 Qy 481 AGAAGGCTTCAGAAAGCAGAGAGAGATTTGTTTATACCAACCTTACATCACTAC 540
 Db 486 AGAAGGCTTCAGAAAGCAGAGAGAGATTTGTTTATACCAACCTTACATCACTAC 545
 Qy 541 TCAAGAGCGTGCATATCCGACGAGGGGGCTGAGAGCTGCGCAAGTGGGGCTTTG 600
 Db 546 TCAAGAGCGTGCATATCCGACGAGGGGGCTGAGAGCTGCGCAAGTGGGGCTTTG 605
 Qy 601 GCATCTCTCATTCGATCCGTGGCTCTCTTCTCACTTCACTGCTTCACTGATTTGAG 660
 Db 606 GCATCTCTCATTCGATCCGTGGCTCTCTTCTCACTTCACTGCTTCACTGATTTGAG 665
 Qy 661 GAATACAGATGGGCGGCCAAATTCGAAACAGCTGATTAAGGTGGAAGATGCAAG 720
 Db 666 GAATACAGATGGGCGGCCAAATTCGAAACAGCTGATTAAGGTGGAAGATGCAAG 725
 Qy 721 ATGATGTCAAGATGGCTTCTCATGAGATCAAAATTTGATTCATTAAGATGGGGCA 780
 Db 726 ATGATGTCAAGATGGCTTCTCATGAGATCAAAATTTGATTCATTAAGATGGGGCA 785
 Qy 781 AAGACCTTACCAAGATGATGATCTTCAACACTGATAGAGATCACTGGAGCAAAATAT 840
 Db 786 AAGACCTTACCAAGATGATGATCTTCAACACTGATAGAGATCACTGGAGCAAAATAT 845
 Qy 841 CCAAGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 Db 846 CCAAGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 905
 Qy 901 ATGATGATGAGCGGTGAGCTTTATATCATGAGAGACCTCTCACTTATTAAGATCTT 960
 Db 906 ATGATGATGAGCGGTGAGCTTTATATCATGAGAGACCTCTCACTTATTAAGATCTT 965
 Qy 961 GGGCTGCTTCAAGAGAGACTCTGCGGTGCTCTGAGATGCAAGAAAGAGTGGGA 1020
 Db 966 GGGCTGCTTCAAGAGAGACTCTGCGGTGCTCTGAGATGCAAGAAAGAGTGGGA 1025

QY 1021 GTTGTCCTTCAGTATTATACAGGTAATATTTTCCACTACAGTCTGGTG 1080
 DB 1026 GTTGTCCTTCAGTATTATACAGGTAATATTTTCCACTACAGTCTGGTG 1085
 QY 1081 ATGAGTCTGACGACGAGAACCTTTTACCCACTGCGCTGCATTTACTGCGAGTAAAG 1140
 DB 1086 ATGAGTCTGACGACGAGAACCTTTTACCCACTGCGCTGCATTTACTGCGAGTAAAG 1145
 QY 1141 GCGAGGCGCATGAGAGAGAGTATAGCTGCTGCGAGCCCTCATATATACACAGTCTG 1200
 DB 1146 GCGAGGCGCATGAGAGAGAGTATAGCTGCTGCGAGCCCTCATATATACACAGTCTG 1205
 QY 1201 CTGAGCCATGAGAGAGAGACAGACATCACTTTTGGATTCAGTGGAGTCTGAGACC 1260
 DB 1206 CTGAGCCATGAGAGAGAGACAGACATCACTTTTGGATTCAGTGGAGTCTGAGACC 1265
 QY 1261 AGTCTACTTGATGACTTATACAGTATTTCTTTCCATCACTCCACGAGACACCATG 1320
 DB 1266 AGTCTACTTGATGACTTATACAGTATTTCTTTCCATCACTCCACGAGACACCATG 1325
 QY 1321 ACTGTCATGATCCAAAGACATGAATGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1380
 DB 1326 ACTGTCATGATCCAAAGACATGAATGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1385
 QY 1381 GTTGTGACAGATGAGAGAAATGCTGCTAGTCC 1416
 DB 1386 GTTGTGACAGATGAGAGAAATGCTGCTAGTCC 1421

RESULT 4
 ID AAV02296 standard; DNA; 1778 BP.
 AAV02296;

XX 21-MAY-1998 (first entry)
 XX Human secreted protein AM282 full-length cDNA clone.
 DE Secreted protein; AM282; cytokine; human; ds.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 17..1435
 FT /*tag= a
 FT sig_peptide 17..88
 FT /*tag= b
 FT mat_peptide 89..1432
 FT /*tag= b
 XX
 XX MO9739030-A2.
 XX 23-OCT-1997.
 XX
 XX 16-APR-1997; 97WO-US06475.
 XX
 XX 13-JAN-1997; 97US-0783520.
 XX 18-APR-1996; 96US-0634325.
 XX
 XX (GEMV) GENETICS INST INC.
 XX
 XX Jacobs K, Lavallie ER, McCoy JM, Werberg D, Racie LA;
 XX Spaulding V;
 XX MPI; 1997-526400/48.
 XX P-PSDB; AAM33604.
 XX
 XX New isolated secretory proteins AM340, AM282 and AK583 - possibly
 XX have cytokine, cell proliferation/differentiation regulating,
 XX immunomodulating activities, etc.

PS Claim 15; Page 44-45; 59pp; English.

XX This cDNA clone encodes a protein (see W33604) designated AM282.
 CC It was identified as "yfp5b10.r1 human EST 30142.5" (GenBank
 CC accession No. R77830) in a database search using a partial AM282
 CC clone (see T97398) obtained from a human foetal kidney cDNA
 CC library using methods selective for cDNAs encoding secreted
 CC proteins. AM282 is deposited in ATCC 98026 together with clones
 CC AM340 (see T97397) and AK583 (see V02297), which are also claimed.
 CC AM282 protein can be used in a claimed method for preventing,
 CC treating or ameliorating a medical condition. It may exhibit
 CC cytokine, cell proliferation (either inducing or inhibiting) or
 CC cell differentiation (either inducing or inhibiting) activity or
 CC may induce production of other cytokines in certain cell
 CC populations. It may also exhibit e.g. immune stimulating or
 CC suppressing activity, haematopoiesis regulating activity, tissue
 CC growth activity, activin/inhibin activity, chemotactic or
 CC chemokinetic activity, haemostatic or thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, tumour
 CC inhibition activity, or other activities. No evidence of any of
 CC these activities is given in the specification.

XX Sequence 1778 BP; 514 A; 386 C; 422 G; 456 T; 0 other;

Query Match 99.8%; Score 1412.8; DB 18; Length 1778;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAATTCCTTATCTTGCGATTTTGGTGGTGTCACTTTATCTCTGCTGGG 60
 DB 17 ATGAATTCCTTATCTTGCGATTTTGGTGGTGTCACTTTATCTCTGCTGGG 76
 QY 61 AAGGCTATATGCAAGATGCGCATCTCAAGAGACTTTGAAGAAATGAAGATA 120
 DB 77 AAGGCTATATGCAAGATGCGCATCTCAAGAGACTTTGAAGAAATGAAGATA 136
 QY 121 GCCAGCTGTGGAGATGTTGCTAAAGCAATCATCACTGTTTATGTTAAAGCCAG 180
 DB 137 GCCAGCTGTGGAGATGTTGCTAAAGCAATCATCACTGTTTATGTTAAAGCCAG 196
 QY 181 AACGATCTATGAGCGATTTGGCACTTCTGTTGATACCTGTTGACCCAGACTGAGTGC 240
 DB 197 AACGATCTATGAGCGATTTGGCACTTCTGTTGATACCTGTTGACCCAGACTGAGTGC 256
 QY 241 TCCAAAGACTTAAAGAAAGCCATCCCAATTAATGACCAAAACCTGACGAAAGTGGCTG 300
 DB 257 TCCAAAGACTTAAAGAAAGCCATCCCAATTAATGACCAAAACCTGACGAAAGTGGCTG 316
 QY 301 GAGAAAGTTCACTGAGCCAGTGAATACCCCACTGGGAGAGGGAGAAAGTCAAGCT 360
 DB 317 GAGAAAGTTCACTGAGCCAGTGAATACCCCACTGGGAGAGGGAGAAAGTCAAGCT 376
 QY 361 GTGATGCTGAGCCAGAAATTCATAGATACCATCTGCGGCTTTGACAGAGCATTTGGG 420
 DB 377 GTGATGCTGAGCCAGAAATTCATAGATACCATCTGCGGCTTTGACAGAGCATTTGGG 436
 QY 421 ACTCTCCAGAAAGCATTAACGACGAGAGTCTGTTGTTGATACCTTTTCATGATCACTGAG 480
 DB 437 ACTCTCCAGAAAGCATTAACGACGAGAGTCTGTTGTTGATACCTTTTCATGATCACTGAG 496
 QY 481 AGAAGGCTCAGAAAGCAAGAGGAGATTTGTTTATATACCAACTTACATCACTAC 540
 DB 497 AGAAGGCTCAGAAAGCAAGAGGAGATTTGTTTATATACCAACTTACATCACTAC 556
 QY 541 TCAAGACGCTGCAATACCGAACGACGAGGCGGTGAGAGCTGCAAGTGGGGCTTTTG 600
 DB 557 TCAAGACGCTGCAATACCGAACGACGAGGCGGTGAGAGCTGCAAGTGGGGCTTTTG 616
 QY 601 GCATCTCTCATTTGATCCGTTGCTCTTCTCCATCTACAGTCTCTCAACAGGATTCAG 660
 DB 617 GCATCTCTCATTTGATCCGTTGCTCTTCTCCATCTACAGTCTCTCAACAGGATTCAG 676
 QY 661 GAATACAGATGCGTCCCAAAATTCAAACAGCTGTATTAACGTTGAGAAATGACAGAA 720

```

Db 677 GAATACAGAGATGGGCTCCAGATTCACAGCTCTATTAGGTGGAGAGACAGAA 736
Qy 721 ATGATGTCAGAGATGGCTTCTCATGGGATCAAAATTTGCTACCTTAAGATGGGGCA 780
Db 737 ATGATGTCAGAGATGGCTTCTCATGGGATCAAAATTTGCTACCTTAAGATGGGGCA 796
Qy 781 AAGACCTACAGAGATGGCTTCTCATGGGATCAAAATTTGCTACCTTAAGATGGGGCA 840
Db 797 AAGACCTACAGAGATGGCTTCTCATGGGATCAAAATTTGCTACCTTAAGATGGGGCA 856
Qy 841 CCAGAACAGGTTGTAAGTGTGATGAGATCTGACAGCTGGGATGTTGGGAGGGTCC 900
Db 857 CCAGAACAGGTTGTAAGTGTGATGAGATCTGACAGCTGGGATGTTGGGAGGGTCC 916
Qy 901 ATGATGTAAGTGGGCTGAGGCTTTATATCATGGGAGACCTCTCATTTAAGATCTT 960
Db 917 ATGATGTAAGTGGGCTGAGGCTTTATATCATGGGAGACCTCTCATTTAAGATCTT 976
Qy 961 GGGGCTGCTCCAGAGAGAGCTGCGGGCTGCTGGAATGCAAGAAACAAGTGA 1020
Db 977 GGGGCTGCTCCAGAGAGAGCTGCGGGCTGCTGGAATGCAAGAAACAAGTGA 1036
Qy 1021 GTTGGTGCCTTCCAGATTTATCATGTTAACAAGTAAATTTCCACTACAGTCTGTC 1080
Db 1037 GTTGGTGCCTTCCAGATTTATCATGTTAACAAGTAAATTTCCACTACAGTCTGTC 1096
Qy 1081 ATGAGTCTGACGAGAGAACTTTTACCCTGCGCTGCAATTACTGCGAGTGAAG 1140
Db 1097 ATGAGTCTGACGAGAGAACTTTTACCCTGCGCTGCAATTACTGCGAGTGAAG 1156
Qy 1141 GCCAGGGCCATCATGAGAGAGGTTATGAGCCCTGCGAGCCCTCAATATCATCTAGTC 1200
Db 1157 GCCAGGGCCATCATGAGAGAGGTTATGAGCCCTGCGAGCCCTCAATATCATCTAGTC 1216
Qy 1201 CTGAGCCATGAGAGAGAGACAGACATCACTTTTGGATCAAGCTGAGTCTGAGACC 1260
Db 1217 CTGAGCCATGAGAGAGAGACAGACATCACTTTTGGATCAAGCTGAGTCTGAGACC 1276
Qy 1261 AGTCTACTTGAATGATTTATCAAGATTTTCTTCCATCACTCCACGAGACACCATG 1320
Db 1277 AGTCTACTTGAATGATTTATCAAGATTTTCTTCCATCACTCCACGAGACACCATG 1336
Qy 1321 ACTGTCATGATTCAGAGAGATGATGTCGTCGCTTGGGCTGTTCTTAT 1380
Db 1337 ACTGTCATGATTCAGAGAGATGATGTCGTCGCTTGGGCTGTTCTTAT 1396
Qy 1381 GTTGTGAGAGATGAGAAATGCTGCTAGTCC 1416
Db 1397 GTTGTGAGAGATGAGAAATGCTGCTAGTCC 1432

```

RESULT 5
AAZ98034
ID AAZ98034 standard; cDNA; 1863 BP.

AC AAZ98034;
DT 09-MAY-2000 (first entry)

DE Human secreted protein encoding nucleotide sequence SEQ ID NO:28.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;
XX tumour; neurodegenerative disorder; developmental abnormality; allergy;
XX foetal deficiency; blood disorder; immune system disorder; arthritis;
XX autoimmune disease; hepatic disease; renal disease; inflammation;
XX Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
XX infection; AIDS; spinal cord injury; transplant rejection; diabetes;
XX asthma; sepsis; acne; psoriasis; cardiovascular disorder;
XX reproductive disorder; gastrointestinal disorder; respiratory disorder;
XX metabolic disorder; food additive; preservative; ss.

```

OS Homo sapiens.
XX
XX MO200004140-A1.
XX
XX 27-JAN-2000.
XX
XX 14-JUL-1999; 99WO-US15849.
XX
XX 15-JUL-1998; 98US-0092921.
XX 15-JUL-1998; 98US-0092922.
XX 15-JUL-1998; 98US-0092956.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
XX Lallier DW, Ehner R, Olsen HS, Brewer LA, Florence KA, Young PE;
XX Mueneski M, Endress GA, Soppet DR;
XX
XX WPI; 2000-161128/14.
XX
XX P-PSDB; AAY87081.
XX
XX New isolated human genes, useful for diagnosis and treatment of, e.g.
XX cancers, neurological or blood disorders
XX
XX Claim 1; Page 319; 494pp; English.
XX
XX The polynucleotide sequences given in AAZ98017 to AAZ98108 encode the
XX human secreted proteins given in AAY87064 to AAY87223. Human secreted
XX protein can have activities based on the tissues and cells the genes are
XX expressed in. Examples of activities include: cytostatic;
XX immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
XX antiallergic. The polynucleotides and their corresponding secreted
XX polypeptides are useful for preventing, treating or ameliorating medical
XX conditions, e.g. by protein or gene therapy. Also pathological conditions
XX can be diagnosed by determining the amount of the new polypeptides in a
XX sample or by determining the presence of mutations in the new
XX polynucleotides. Human secreted proteins and their polynucleotides can
XX be used for developing products for the diagnosis or treatment of cancer,
XX tumours, neurodegenerative disorders, developmental abnormalities and
XX foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, hepatic and renal disease, inflammation,
XX allergies, Alzheimer's disease, behavioural disorder, schizophrenia,
XX osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
XX cardiovascular disorders, reproductive disorders, gastrointestinal
XX disorders, respiratory disorders and metabolic disorders. The
XX proteins or polynucleotides can also be used as food additives or
XX preservatives. The proteins are also useful for identifying their
XX binding partners. AAZ98008 to AAZ98016 and AAY87063 are sequence used in
XX the exemplification of the present invention.
XX
XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

```

Query Match 99.8%; Score 1412.8; DB 21; Length 1863;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 ATGAAATTCCTTATCTTGGATTTTGGGCTGCTTACCTTTATCCCTGCTGGG 60
Db 99 ATGAAATTCCTTATCTTGGATTTTGGGCTGCTTACCTTTATCCCTGCTGGG 158
Qy 61 AAAGCTATATGAGAGATGGCATCTTAAGAGAGCTTTTGAAGAAATTAAGAGAAATA 120
Db 159 AAAGCTATATGAGAGATGGCATCTTAAGAGAGCTTTTGAAGAAATTAAGAGAAATA 218
Qy 121 GCCAGCTGTGAGAGATGTTGCTTAAGACATCATCACTAGCTGTTATGTTAAAGCCAG 180
Db 219 GCCAGCTGTGAGAGATGTTGCTTAAGACATCATCACTAGCTGTTATGTTAAAGCCAG 278
Qy 181 AACGATCTTATGAGAGATGTTGCTTGTGTTATCTGTTGAGCCAGCTAGTGGC 240
Db 279 AACGATCTTATGAGAGATGTTGCTTGTGTTATCTGTTGAGCCAGCTAGTGGC 338

```


QY 241 TCAGAGACCTAGAAAAAGCCATCCAAATTATGTACCAAAACTGACAGAGATGGGCTG 300
 Db 339 TCAGAGACCTAGAAAAAGCCATCCAAATTATGTATCAAAAAGCTGACAGAGATGGGCTG 398
 QY 301 GAGAAAGTTCACTGAGAGCCAGTGAATACCCACTGGGAGAGGGAGAAATCAAGCT 360
 Db 399 GAGAAAGTTCACTGAGAGCCAGTGAATACCCACTGGGAGAGGGAGAAATCAAGCT 458
 QY 361 GTGATGCTGGAGCCAGAAATCATTAAGTAGCAATCCGAGCTGGGAGAGAGATGGG 420
 Db 459 GTGATGCTGGAGCCAGAAATCATTAAGTAGCAATCCGAGCTGGGAGAGAGATGGG 518
 QY 421 ACTCTCTCAGAAAGCATTACAGCAAAAGTTCTGTGTGATCACTTTTCATGAATGCTGAG 480
 Db 519 ACTCTCTCAGAAAGCATTACAGCAAAAGTTCTGTGTGATCACTTTTCATGAATGCTGAG 578
 QY 481 AGAAGGCTCTCAGAAAGCAGAGAGAAATTTGTTTATTAACAACCTTACATCACTAC 540
 Db 579 AGAAGGCTCTCAGAAAGCAGAGAGAAATTTGTTTATTAACAACCTTACATCACTAC 638
 QY 541 TCAGAGACGCTGCAATCCGAAAGCAGAGGGGCTGGAGAGCTGCAAGGTGGGGCTTTG 600
 Db 639 TCAGAGACGCTGCAATCCGAAAGCAGAGGGGCTGGAGAGCTGCAAGGTGGGGCTTTG 698
 QY 601 GCATCTCTCATTCGATCGTGGCTCTCTCTCTCAATCACTACAGTCTCAACAGTATTCAG 660
 Db 699 GCATCTCTCATTCGATCGTGGCTCTCTCTCTCAATCACTACAGTCTCAACAGTATTCAG 758
 QY 661 GAATACCAAGATGCGTGGCTCCCAAAATTCGAACAGCTGTATTAACGTTGGAAGATGCAGAA 720
 Db 759 GAATACCAAGATGCGTGGCTCCCAAAATTCGAACAGCTGTATTAACGTTGGAAGATGCAGAA 818
 QY 721 ATGATGTCAAAAGTGGCTCTCATGAGGATCAAAATGTGATTCATTCAGCTTAAAGTGGGGCA 780
 Db 819 ATGATGTCAAAAGTGGCTCTCATGAGGATCAAAATGTGATTCATTCAGCTTAAAGTGGGGCA 878
 QY 781 AAGACCTAACCAAGATGATGATTCCTTCAACACTGAGAGAGATCACTGGAGCAAAATAT 840
 Db 879 AAGACCTAACCAAGATGATGATTCCTTCAACACTGAGAGAGATCACTGGAGCAAAATAT 938
 QY 841 CCAGAACAGGTTGTACTGTGTGATGTGACATCTGCAACAGCTGGAGTGGGCAAGGTGCC 900
 Db 939 CCAGAACAGGTTGTACTGTGTGATGTGACATCTGCAACAGCTGGAGTGGGCAAGGTGCC 998
 QY 901 ATGATGATGAGCGGTGGAGCCTTTATATCATGGAAGACACTCACTTATTAAGATCTT 960
 Db 999 ATGATGATGAGCGGTGGAGCCTTTATATCATGGAAGACACTCACTTATTAAGATCTT 1058
 QY 961 GGGCTGCGTCCAAAGAGAGACTCTGGGCTGGTCTGGAATGCAAGAGAAAGAGTGA 1020
 Db 1059 GGGCTGCGTCCAAAGAGAGACTCTGGGCTGGTCTGGAATGCAAGAGAAAGAGTGA 1118
 QY 1021 GTTGTGTGCTTCAAGTATTTATCAAGTAAATATTTTCCAACTACAGTCTGTG 1080
 Db 1119 GTTGTGTGCTTCAAGTATTTATCAAGTAAATATTTTCCAACTACAGTCTGTG 1178
 QY 1081 ATGAGATGTGACGAGAGAACTTCTTACCACTGGGCTGCAATTAATGAGAGAAAG 1140
 Db 1179 ATGAGATGTGACGAGAGAACTTCTTACCACTGGGCTGCAATTAATGAGAGAAAG 1238
 QY 1141 GCCAGGGCCATCATGAGAGAGGTTATGAGCTGTGAGCCCTCAATATATCATCAAGTCTC 1200
 Db 1239 GCCAGGGCCATCATGAGAGAGGTTATGAGCTGTGAGCCCTCAATATATCATCAAGTCTC 1298
 QY 1201 CTGAGCCATGAGAGAGAGAGACATCAACTTTTGGATCCAAGCTGGAGTGGAGAGCC 1260
 Db 1299 CTGAGCCATGAGAGAGAGAGACATCAACTTTTGGATCCAAGCTGGAGTGGAGAGCC 1358
 QY 1261 AGTCTACTTGAATGATTTATCAAGATTTTCTTCTTCCATCACTCCCAAGAGAGACCACTG 1320
 Db 1359 AGTCTACTTGAATGATTTATCAAGATTTTCTTCTTCCATCACTCCCAAGAGAGACCACTG 1418
 QY 1321 ACTGTATGATTCAGAAAGCAGATGAATGTGCTGTGTTGGGCTGTTGTTCTTAT 1380

Db 1419 ACTGTCTGATCCAAAGCAGATGAATGTGCTGCTGCTGTGGCTGTGTTCTTAT 1478
 QY 1381 GTTCTTGACAGATGAGAAAGATGCTGCTGAGTCC 1416
 Db 1479 GTTCTTGACAGATGAGAAAGATGCTGCTGAGTCC 1514
 RESULT 6
 AAD11647
 ID AAD11647 standard; cDNA; 1863 BP.
 XX
 AC AAD11647;
 XX
 DT 24-SEP-2001 (first entry)
 XX
 DE Human secreted protein-encoding gene 18 cDNA clone HRACT35, SEQ ID NO:28.
 KW Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KW cell culture; chemotaxis; vulnery; binding partner identification;
 KW gene therapy; ss.
 KW Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT 99..1517
 FT CDS /tag= a
 FT /product= "Human secreted protein precursor"
 FT sig_peptide 99..170
 FT /tag= b
 FT mat_peptide 171..1514
 FT /tag= c
 FT /product= "Mature human secreted protein"
 FT
 PN MO200151504-A1.
 PN
 XX
 PD 19-JUL-2001.
 PD
 XX
 PF 12-JAN-2001; 2001W0-US00911.
 PF
 XX
 PR 13-JAN-2000; 2000US-0482273.
 PR
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 PI Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DM, Olesen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Endress GA, Muscenski M, Ebner R;
 XX
 DR WPI: 2001-425865/45.
 DR P-PSDB; AAB06058.
 DR
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 PT
 XX
 PS Claim 1; Page 686; 86app; English.
 PS
 CC AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
 CC protein genes, and AAB06041-AAB06132 represent the proteins they encode.
 CC AAB06133-AAB06205 represent human secreted protein fragments.
 CC The secreted proteins and their genes are useful for preventing, treating
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 CC Pathological conditions can be diagnosed, e.g., by determining the amount of the
 CC new protein in a sample or by determining the presence of mutations in
 CC the new genes. Specific uses are described for each of the 71 genes,
 CC based on the tissues in which they are most highly expressed, and include
 CC developing products for the diagnosis or treatment of proliferative

disorders, cancer, tumours, foetal and developmental abnormalities,
 haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 angiotensin disorders, kidney disorders, gastrointestinal disorders,
 pregnancy-related disorders, endocrine disorders, and infections. The
 proteins can also be used to aid wound healing and epithelial cell
 proliferation, to prevent skin aging due to sunburn, to maintain organs
 before transplantation, for supporting cell culture of primary tissues,
 to regenerate tissues, to identify their cognate ligands or binding
 partners, and in chemotaxis, and can be used as a food additive or
 preservative to modify storage properties. Antibodies specific for a
 protein of the invention can be used in alleviating symptoms associated
 with the disorders mentioned above, and in diagnostic immunoassays e.g.,
 radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
 CC The present sequence represents a human secreted protein-coding cDNA of
 CC the invention.

XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 99.8%; Score 1412.8; DB 22; Length 1863;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAATCTTATCTTGGACATTTTGGTGGTCTTACCTTTATCCCTGCTCTGGG 60
 99 ATGAATCTTATCTTGGACATTTTGGTGGTCTTACCTTTATCCCTGCTCTGGG 158
 61 AAAGCTATATGCAAGATGAGCATCTCTAAGAGACATTTTGAAGAAATAAAGAAATA 120
 159 AAAGCTATATGCAAGATGAGCATCTCTAAGAGACATTTTGAAGAAATAAAGAAATA 218
 121 GCCAGCTGTGAGATGTTGCTTAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 180
 219 GCCAGCTGTGAGATGTTGCTTAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 278
 181 AACGATCTTATGACGAGATGCTTGTGTTATGTTGAGCCAGACAGCTGAGTGGC 240
 279 AACGATCTTATGACGAGATGCTTGTGTTATGTTGAGCCAGACAGCTGAGTGGC 338
 241 TCCAAAGACTTGAAGAAAGCCATCCAAATTTATGTAACAAACCTGACAGAGATGGCTG 300
 339 TCCAAAGACTTGAAGAAAGCCATCCAAATTTATGTAACAAACCTGACAGAGATGGCTG 398
 301 GAGAAAGTTCACTGTGAGCCAGTGAATATCCCACTGGAGAGGGGAGAAAGATCACT 360
 399 GAGAAAGTTCACTGTGAGCCAGTGAATATCCCACTGGAGAGGGGAGAAAGATCACT 458
 361 GTGATGCTGAGAGCCAGAAATTCATTAAGTATGCTTGGCTTGGAGACAGCATTTGGG 420
 459 GTGATGCTGAGAGCCAGAAATTCATTAAGTATGCTTGGCTTGGAGACAGCATTTGGG 518
 421 ACTCTTCCAGAGGAGCATTAACAGACAGAGTTCTGGTGGTGAACCTCTTTGATGAATCGAG 480
 519 ACTCTTCCAGAGGAGCATTAACAGACAGAGTTCTGGTGGTGAACCTCTTTGATGAATCGAG 578
 481 AGAAGGAGCTTCAAGAGCAAGAGAGAGATGTTGTTTAAACCACTTATCAATCACTAC 540
 579 AGAAGGAGCTTCAAGAGCAAGAGAGAGATGTTGTTTAAACCACTTATCAATCACTAC 638
 541 TCAAGAGAGGTCATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 639 TCAAGAGAGGTCATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 698
 601 GCATCTCTCATTCGATCCGTCGCTCTCTTCCATCTCAAGTCCCTACAGAGATTTAG 660
 699 GCATCTCTCATTCGATCCGTCGCTCTCTTCCATCTCAAGTCCCTACAGAGATTTAG 758
 661 GAATACAGAGATGCGGTGCCCAAAATTCACAGAGCTGTATTAAGTGAAGATGCAGAA 720
 759 GAATACAGAGATGCGGTGCCCAAAATTCACAGAGCTGTATTAAGTGAAGATGCAGAA 818

721 ATGATGTCAGAGATGCTTCTCATGGAGTCAAAATTTCTATTCAGTAAAGATGGGGCA 780
 819 ATGATGTCAGAGATGCTTCTCATGGAGTCAAAATTTCTATTCAGTAAAGATGGGGCA 878
 781 AAGACTTACCAGATGATGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAATAT 840
 879 AAGACTTACCAGATGATGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAATAT 938
 841 CCAGAACAGTTGATCTGTGTCAGTGCATCTGACAGCTGGAGATTTGGGACAGGTGCC 900
 939 CCAGAACAGTTGATCTGTGTCAGTGCATCTGACAGCTGGAGATTTGGGACAGGTGCC 998
 901 ATGATGATGAGCGGTGAGAGCTTTATATCATGAGGAGACATCTCACTTATTAAGATCTT 960
 999 ATGATGATGAGCGGTGAGAGCTTTATATCATGAGGAGACATCTCACTTATTAAGATCTT 1058
 961 GGGCTGCGTCCAAAGAGAGACTGTGCGGTGCTGTGACCTGACAGAAACAAGTGA 1020
 1059 GGGCTGCGTCCAAAGAGAGACTGTGCGGTGCTGTGACCTGACAGAAACAAGTGA 1118
 1021 GTTGTGCTTCCAGATTTATCACTTACAGAGTAAATTTTTCATCACTACAGTCTGTG 1080
 1119 GTTGTGCTTCCAGATTTATCACTTACAGAGTAAATTTTTCATCACTACAGTCTGTG 1178
 1081 ATGAGATCTGACGAGAGACCTTCTTACCCAGCTGGGCTGCAATTCACCTGCAAGTGA 1140
 1179 ATGAGATCTGACGAGAGACCTTCTTACCCAGCTGGGCTGCAATTCACCTGCAAGTGA 1238
 1141 GCCAGGCGCATCATGAGAGAGATTTAGAGCTGTGAGAGCCCTCAATATCACTAGTGC 1200
 1239 GCCAGGCGCATCATGAGAGAGATTTAGAGCTGTGAGAGCCCTCAATATCACTAGTGC 1298
 1201 CTGAGCCATGAGAGAGAGAGAGAGATTTTGTGATCCAGCTGAGAGGCTTGGAGCC 1260
 1299 CTGAGCCATGAGAGAGAGAGAGAGATTTTGTGATCCAGCTGAGAGGCTTGGAGCC 1358
 1261 AGTCTACTGATGACTTATCAAGATTTTCTTCTTCCATCACTGCCAGAGACACCATG 1320
 1359 AGTCTACTGATGACTTATCAAGATTTTCTTCTTCCATCACTGCCAGAGACACCATG 1418
 1321 ACTGTCTGATGATCCAAAGAGAGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTAT 1380
 1419 ACTGTCTGATGATCCAAAGAGAGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTAT 1478
 1381 GTTGTGCAAGATGAGAGAAATGCTGCTAGTCC 1416
 1479 GTTGTGCAAGATGAGAGAAATGCTGCTAGTCC 1514

RESULT 7
 ABR69743
 ID ABR69743 standard; cDNA; 1863 BP.
 AC ABR69743;
 DT 15-JUL-2002 (first entry)
 XX
 XX Human secreted protein gene 18 #1.
 DE
 DE Human; ss; gene; secreted protein; gene therapy; immunosuppressive;
 KW antiatheritic; antirheumatic; antiproliferative; cytostatic; antibacterial;
 KW vasoactive; cerebroprotective; neurotropic; neuroprotective; neoplasm;
 KW viruslike; fungicide; ophthalmological; autoimmune disease; cardiac arrest;
 KW rheumatoid arthritis; hyperproliferative disorder; cerebral ischemia;
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischemia;
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
 KW ocular disorder; corneal infection; wound healing; skin aging;
 KW epithelial cell proliferation; food additive.
 OS Homo sapiens.
 XX
 XX
 PN W0200226931-A2.

XX 04-APR-2002.
 PD 24-SEP-2001; 2001WO-US29871.
 XX 25-SEP-2000; 2000US-234925P.
 PR 12-JAN-2001; 2001WO-US00911.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Komatsu J G, Duan DR, Rosen CA, Moore PA, Shi Y,
 PI Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Endress GA, Mucenki M, Ebner R,
 XX WPI; 2002-362489/39.
 DR P-PSDB; ABG33880.
 XX Novel 71 isolated secreted polypeptides and polynucleotides encoding
 PT the polypeptides, useful for treating Huntington's disease, sepsis,
 PT meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
 PT asthma
 XX
 XX Claim 1; Page 1170; 1478pp; English.
 XX The invention relates to an isolated nucleic acid molecule (or its
 CC fragment, homologue complement or allelic variant) encoding a human
 CC secreted protein (and its fragment, domain, epitope, variant, secreted
 CC form and species variant). Also included are a recombinant vector
 CC comprising the nucleic acid, a recombinant host cell comprising the
 CC vector, an antibody against the secreted protein, a recombinant host cell
 CC that expresses the secreted protein and a method of identifying a binding
 CC partner of the secreted protein. The nucleic acid and protein are used to
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
 CC for example autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiodenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
 CC ocular disorders e.g. corneal infection. Many other diseases and
 CC disorders are listed in the specification. The polypeptides can also be
 CC used to aid wound healing an epithelial cell proliferation, to prevent
 CC skin aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities. The present
 CC sequence encodes a novel human secreted protein of the invention.
 CC
 XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
 SO
 Query Match 99.8%; Score 1412.8; DB 24; Length 1863;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 301 GAGAAAGTCACTGAGCAGTGAATACCCCACTGGGAGAGAGGAAGATCAGCT 360
 DB 399 GAGAAAGTCACTGAGCAGTGAATACCCCACTGGGAGAGAGGAAGATCAGCT 458
 QY 361 GTGATGCTGAGCCCAAGATTCATAGATAGCCATCTGGGCTTGGCAGCAGATTGGG 420
 DB 459 GTGATGCTGAGCCCAAGATTCATAGATAGCCATCTGGGCTTGGCAGCAGATTGGG 518
 QY 421 ACTCTCCAGAGGCAATTAACGACGAGAGTTCTGGTGTGACCTTTTCGATGAATCTGAG 480
 DB 519 ACTCTCCAGAGGCAATTAACGACGAGAGTTCTGGTGTGACCTTTTCGATGAATCTGAG 578
 QY 481 AGAAGGCGCTCAGAGGCAAGGAGGGAATTTGTTTATATACCAACTTACATCACTAC 540
 DB 579 AGAAGGCGCTCAGAGGCAAGGAGGGAATTTGTTTATATACCAACTTACATCACTAC 638
 QY 541 TCAGAGCAGGTGCAATACCGACGAGGCGGTGAGAGCTGCCAAGTGGGCGCTTTTG 600
 DB 639 TCAGAGCAGGTGCAATACCGACGAGGCGGTGAGAGCTGCCAAGTGGGCGCTTTTG 698
 QY 601 GCATCTCATTCGATCGTGCGCTCTTCTCCATCTACAGTCTCAACAGGATTTTCAG 660
 DB 699 GCATCTCATTCGATCGTGCGCTCTTCTCCATCTACAGTCTCAACAGGATTTTCAG 758
 QY 661 GAATACAGAGTGGCGGCGCCCAAAATTCGAACGCGCTGTTTACGTTGGAAGATGACGA 720
 DB 759 GAATACAGAGTGGCGGCGCCCAAAATTCGAACGCGCTGTTTACGTTGGAAGATGACGA 818
 QY 721 ATGATGTCAGAAATGCGCTTCATGATGAGATCAAAATTTGATTCAGTAAAGATGGGGCA 780
 DB 819 ATGATGTCAGAAATGCGCTTCATGATGAGATCAAAATTTGATTCAGTAAAGATGGGGCA 878
 QY 781 AAGACTTACCCAGATTAATCTTCTTCAACATGTAAGAGATCACTGGAGCAAAATAT 840
 DB 879 AAGACTTACCCAGATTAATCTTCTTCAACATGTAAGAGATCACTGGAGCAAAATAT 938
 QY 841 CCAGAACAGGTTGATCTGTCAGTGAATCTGAGCAGCTGGAGATGTTGGCAGAGGGTGC 900
 DB 939 CCAGAACAGGTTGATCTGTCAGTGAATCTGAGCAGCTGGAGATGTTGGCAGAGGGTGC 998
 QY 901 ATGATGATGCGCGGTGAGCGCTTATATCATGAGGAGACATCTCACTTATTAAGATCTT 960
 DB 999 ATGATGATGCGCGGTGAGCGCTTATATCATGAGGAGACATCTCACTTATTAAGATCTT 1058
 QY 961 GGGCTGCTCCAAAGAGAGACTCTGGCGTGTGCTCTGACCTGACGAGAAACAAGTGA 1020
 DB 1059 GGGCTGCTCCAAAGAGAGACTCTGGCGTGTGCTCTGACCTGACGAGAAACAAGTGA 1118
 QY 1021 GTTGTGCGCTTCCAGATTAATCAATCAATCAAGATTAATTTTCAACTACATCTGCTGTG 1080
 DB 1119 GTTGTGCGCTTCCAGATTAATCAATCAATCAAGATTAATTTTCAACTACATCTGCTGTG 1178
 QY 1081 ATGAGTCTGACGAGAGAACCTTCTTACCACTGGCGTGCAGTAATCACTGGCAGTGAAG 1140
 DB 1179 ATGAGTCTGACGAGAGAACCTTCTTACCACTGGCGTGCAGTAATCACTGGCAGTGAAG 1238
 QY 1141 GCCAGGGCCATCATGAGAGAGGTTATGAGCTGTGACGCCCTCAATATCATCTACAGTGC 1200
 DB 1239 GCCAGGGCCATCATGAGAGAGGTTATGAGCTGTGACGCCCTCAATATCATCTACAGTGC 1298
 QY 1201 CTGAGCCATGAGAGAGGACAGACATCACTTTTGAATCCAGTGTGAGAGTCTGAGGCC 1260
 DB 1299 CTGAGCCATGAGAGAGGACAGACATCACTTTTGAATCCAGTGTGAGAGTCTGAGGCC 1358
 QY 1261 AGTCTACTGATGATTAATCAAGTATTTCTTTCATCATCACTGCCAGGAGACACATG 1320
 DB 1359 AGTCTACTGATGATTAATCAAGTATTTCTTTCATCATCACTGCCAGGAGACACATG 1418
 QY 1321 ACTGTCAATGATCCAAAGCAGATGAATGTTGCTGCTGTGCTGTTGGCTGTTTCTTAT 1380
 DB 1419 ACTGTCAATGATCCAAAGCAGATGAATGTTGCTGCTGTGCTGTTGGCTGTTTCTTAT 1478

Qy 1381 GTGTTCGACATGAGAAATGCTGCTAGTCC 1416
Db 1479 GTGTTCGACATGAGAAATGCTGCTAGTCC 1514

RESULT 8

ACC50817
ID ACC50817 standard; cDNA; 1863 BP.

ACCS0817;

12-JUN-2003 (first entry)

Human secreted protein coding sequence, SEQ ID 484.

Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
vulnerary; antiinflammatory; nootropic; neuroprotective;
antiparkinsonian; gene therapy; human; cardiovascular disorder;
gene; ss.

Homo sapiens.

MO200295010-A2.

28-NOV-2002.

19-MAR-2002; 2002MO-US09785.

21-MAR-2001; 2001US-277340P.

19-JUL-2001; 2001US-306171P.

13-NOV-2001; 2001US-331287P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2003-129429/12.

Novel human secreted proteins, useful for detecting, preventing,
diagnosing, prognosticating, treating and/or ameliorating
cardiovascular disorders such as arrhythmia -

Claim 21; SEQ ID 484; 1881bp; English.

The present invention relates to novel human secreted proteins
(ABR4733-ABR48145) and their coding sequences (ACC50344-ACC50856). The
proteins and their coding sequences are useful for the preparation of a
diagnostic or pharmaceutical composition for diagnosing or treating a
cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
coronary arteriosclerosis and myocardial ischemia), neural disorders,
immune system disorders, muscular disorders, reproductive disorders,
gastrointestinal disorders, pulmonary disorders, renal disorders,
proliferative disorders and/or cancerous diseases and conditions, for
wound healing and epithelial cell proliferation, to treat inflammation or
infection, for treating thrombosis and arteriosclerosis, for treating or
preventing neural damage which occurs in neuronal disorders or
neurodegenerative conditions such as Alzheimer's disease and Parkinson's
disease, to enhance bone and periodontal regeneration and aid in tissue
transplants or bone grafts, to prevent skin aging or hair loss, to
stimulate growth and differentiation of haematopoietic cells and bone
marrow cells when used in combination with other cytokines, to maintain
organs before transplantation or for supporting cell culture of primary
tissues, to increase or decrease differentiation or proliferation of
embryonic stem cells, or to modulate mammalian characteristics or
metabolism.
Note: The sequence data for this patent was published in electronic
format and is available from WIPO at
ftp.wipo.int/pub/published_pct_sequences.

Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 99.8%; Score 1412.8; DB 25; Length 1863;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAATTCCTTACTTCGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 60
Db 99 ATGAATTCCTTACTTCGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 158
Qy 61 AAGCTATATGCAAGAAATGGCATCTTAAGAGACCTTTGAGAAATGAAGAAATA 120
Db 159 AAGCTATATGCAAGAAATGGCATCTTAAGAGACCTTTGAGAAATGAAGAAATA 218
Qy 121 GCCAGCTGTGAGAAATGTTGCTTAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 180
Db 219 GCCAGCTGTGAGAAATGTTGCTTAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 278
Qy 181 AACGATCTTATGAGCGATTTGGCAGCTCTGGTTGATCTGTGAGACCGAGCTAGTGGC 240
Db 279 AACGATCTTATGAGCGATTTGGCAGCTCTGGTTGATCTGTGAGACCGAGCTAGTGGC 338
Qy 241 TCCAAAGACCTGAGAAAGCCATCCAAATTAATGTAACCAAACTGACAGAAATGGGCTG 300
Db 339 TCCAAAGACCTGAGAAAGCCATCCAAATTAATGTAACCAAACTGACAGAAATGGGCTG 398
Qy 301 GAGAAAGTTCACTGAGAGCAATGAGAAATCCCACTGGAGAGGGAGAAATCAAGT 360
Db 399 GAGAAAGTTCACTGAGAGCAATGAGAAATCCCACTGGAGAGGGAGAAATCAAGT 458
Qy 361 GTGATGCTGAGAGCAATGAGAAATCCCACTGGAGAGGGAGAAATCAAGT 420
Db 459 GTGATGCTGAGAGCAATGAGAAATCCCACTGGAGAGGGAGAAATCAAGT 518
Qy 421 ACTCCTCAGAGAGCATTTAAGCAGAAAGTTCTGGTGTGACCTTTGATGAATCTGAG 480
Db 519 ACTCCTCAGAGAGCATTTAAGCAGAAAGTTCTGGTGTGACCTTTGATGAATCTGAG 578
Qy 481 AGAAGGCTTCAGAGCAAGAGAGAGATGTTGTTATTAACCACTTACATCACTAC 540
Db 579 AGAAGGCTTCAGAGCAAGAGAGAGATGTTGTTATTAACCACTTACATCACTAC 638
Qy 541 TCAAGAGCGGTCATTCAG 600
Db 639 TCAAGAGCGGTCATTCAG 698
Qy 601 GCATCTCTCATTCGATCGTGGCTCTCTCTCATCTCACTCACTCACTCACTCACTCACT 660
Db 699 GCATCTCTCATTCGATCGTGGCTCTCTCTCATCTCACTCACTCACTCACTCACTCACT 758
Qy 661 GAATACAGAGATGGCGTGGCCCAAAATTCACACAGCTGTATTCAGTGGAGATGCAAG 720
Db 759 GAATACAGAGATGGCGTGGCCCAAAATTCACACAGCTGTATTCAGTGGAGATGCAAG 818
Qy 721 ATGATGTCAGAAATGGCTTCTCATGAGATCAAAATGTCATTCAGTGAAGATGGGGCA 780
Db 819 ATGATGTCAGAAATGGCTTCTCATGAGATCAAAATGTCATTCAGTGAAGATGGGGCA 878
Qy 781 AAGACTTACCAAGATTCATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAAT 840
Db 879 AAGACTTACCAAGATTCATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAAT 938
Qy 841 CCAGAACAGGTTGATGTCAGTGAATCTGACAGCTGGAGATGTTGGGACAGGTGCC 900
Db 939 CCAGAACAGGTTGATGTCAGTGAATCTGACAGCTGGAGATGTTGGGACAGGTGCC 998
Qy 901 ATGATGATGGCGGTGAGAGCTTTATATCATGGAAGACATCTTCACTTATTAAGATCT 960
Db 999 ATGATGATGGCGGTGAGAGCTTTATATCATGGAAGACATCTTCACTTATTAAGATCT 1058
Qy 961 GGGCTGCTCCAAAGAGAGACTCTGGGCTGTGCTGAGCTGAGAGAGAGAGAGAGAGTGA 1020
Db 1059 GGGCTGCTCCAAAGAGAGACTCTGGGCTGTGCTGAGCTGAGAGAGAGAGAGAGTGA 1118
Qy 1021 GTTGTGCTTCCTCAGATTAATCAAGTTACACAGGTAATATTTTCAACTACAGTCTGGT 1080
Db 1119 GTTGTGCTTCCTCAGATTAATCAAGTTACACAGGTAATATTTTCAACTACAGTCTGGT 1178

QY 1081 ATGAGTCTGACGAGGAACTCTTACCACTGGGCTGCAATTCACTGCGATGAAAG 1140
 Db 1179 ATGAGTCTGACGAGGAACTCTTACCACTGGGCTGCAATTCACTGCGATGAAAG 1238
 QY 1141 GCCAGGGCCATCATGAGGAGGTTATGAGCCTGTGACGCCCTCAATATGACTGAGTC 1200
 Db 1239 GCCAGGGCCATCATGAGGAGGTTATGAGCCTGTGACGCCCTCAATATGACTGAGTC 1298
 QY 1201 CTGAGCCATGAGGAGGACAGCATCACTTTTGGATCCAACTGAGTGCCTGAGACC 1260
 Db 1299 CTGAGCCATGAGGAGGACAGCATCACTTTTGGATCCAACTGAGTGCCTGAGACC 1358
 QY 1261 AGTCTACTTGATGACTTATACAGATATTTCTTCCATCACTCCACGAGACACCATG 1320
 Db 1359 AGTCTACTTGATGACTTATACAGATATTTCTTCCATCACTCCACGAGACACCATG 1418
 QY 1321 ACTGTCATGATCCAAAGCAGATGATGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1380
 Db 1419 ACTGTCATGATCCAAAGCAGATGATGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1478
 QY 1381 GTTGTGACAGCATGAGGAAATGCTGCTAGTCC 1416
 Db 1479 GTTGTGACAGCATGAGGAAATGCTGCTAGTCC 1514
 RESULT 9
 ABZ71453
 ID ABZ71453 standard; cDNA, 1863 BP.
 AC ABZ71453;
 XX
 DT 04-APR-2003 (first entry)
 DE Secreted protein-encoding gene 142 cDNA clone HRACJ35, SEQ ID NO:274.
 XX
 KM Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KM mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KM biliary tract; pancreas; cancer; tumor; hyperproliferative disorder;
 KM immune disorder; inflammation; infection; wound healing; drug screening;
 KM chromosome identification; chromosome mapping; cytostatic; gene therapy;
 KM antiinflammatory; immunosuppressive; veterinary; chromosome 8q22.2;
 KM gene; ss.
 XX
 OS Homo sapiens.
 PN WO200276488-A1.
 PD 03-OCT-2002.
 PF 19-MAR-2002; 2002WO-US08276.
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 DR WPI; 2003-029900/02.
 DR P-PSDB; ABR00274.
 PT New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 PT e.g. gastrointestinal diseases and disorders, or cancers -
 XX
 PS Claim 21; Page 918; 1216pp; English.
 CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,

CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein-
 CC encoding cDNA clone of the invention.
 XX
 SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
 Query Match 99.8%; Score 1412.8; DB 25; Length 1863;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGAAATTCCTTATCTTGGCATTTTGGTGTGTTCACTTATCCCTGTGCTGGG 60
 Db 99 ATGAAATTCCTTATCTTGGCATTTTGGTGTGTTCACTTATCCCTGTGCTGGG 158
 QY 61 AAAGCTATATGCAAGAAATGCGATCTTAAAGAGACCTTTGAAAGAAATAAAGAAATA 120
 Db 159 AAAGCTATATGCAAGAAATGCGATCTTAAAGAGACCTTTGAAAGAAATAAAGAAATA 218
 QY 121 GCCAGCTGTGAGATGTTGTGTAAGCAATCACTAGCTGTTATGTTAAAGCCAG 180
 Db 219 GCCAGCTGTGAGATGTTGTGTAAGCAATCACTAGCTGTTATGTTAAAGCCAG 278
 QY 181 AACGATCTATGAGCGATTTGGCACTTCTGTTGATCTGTGAGCCAGACTGAGTGGC 240
 Db 279 AACGATCTATGAGCGATTTGGCACTTCTGTTGATCTGTGAGCCAGACTGAGTGGC 338
 QY 241 TCCAAGAACTTAAGAAAAGCATCCAAATTAATGTAACCAAACTGCGAAGATGGCTG 300
 Db 339 TCCAAGAACTTAAGAAAAGCATCCAAATTAATGTAACCAAACTGCGAAGATGGCTG 398
 QY 301 GAGAAATTCACCTGAGGCCAGTGAATACCCACCTGGAGAGGGAGAAATCAAGCT 360
 Db 399 GAGAAATTCACCTGAGGCCAGTGAATACCCACCTGGAGAGGGAGAAATCAAGCT 458
 QY 361 GTGATGCTGAGGCCAAGATTTCAATAGATAGCCATCTGGGCTTGGCAGCAGCATTTGG 420
 Db 459 GTGATGCTGAGGCCAAGATTTCAATAGATAGCCATCTGGGCTTGGCAGCAGCATTTGG 518
 QY 421 ACTCTCCAGAGGCAATTAACAGCAAGATTTCTGTTGATCACTTTTGATGAATCTGAG 480
 Db 519 ACTCTCCAGAGGCAATTAACAGCAAGATTTCTGTTGATCACTTTTGATGAATCTGAG 578
 QY 481 AGAAGGCTCCAGAGGAAGAGGAATGTTTATTAACCAACTTATCACTTAACATAC 540
 Db 579 AGAAGGCTCCAGAGGAAGAGGAATGTTTATTAACCAACTTATCACTTAACATAC 638
 QY 541 TCAAGACAGGTGCAATACCAACGAGGCGGTGGAAGCTGCAAGTGGGGCTTTTG 600
 Db 639 TCAAGACAGGTGCAATACCAACGAGGCGGTGGAAGCTGCAAGTGGGGCTTTTG 698
 QY 601 GCATCTCTCATTTGATCGTGGTCTCTCTTCTCACTTCAAGTCTTCAACAGATTTAG 660
 Db 699 GCATCTCTCATTTGATCGTGGTCTCTCTTCTCACTTCAAGTCTTCAACAGATTTAG 758
 QY 661 GAATACAGAGATGGCGGCCAAATTCGAAGCCGTATTAGTGGGAAGATGCGAA 720
 Db 759 GAATACAGAGATGGCGGCCAAATTCGAAGCCGTATTAGTGGGAAGATGCGAA 818
 QY 721 ATGATGTCAGAAATGGCTTCTCATGGGATCAAAATTTGATTCAGTAAAGATGGGGCA 780

```

Db      819 ATGATGTCAGAGATGCTTCTTCATGGAGTCAAAATGTTCATTCAGCTAAAGTGGGGCA 878
Qy      781 AAGACTACCCAGATGATGATCTTCTTCAACACTAGAGAGATGATCTGGGAGCAAAAT 840
Db      879 AAGACTACCCAGATGATGATCTTCTTCAACACTAGAGAGATGATCTGGGAGCAAAAT 938
Qy      841 CCAGAACAGCTTGTACTGCTGATGAGATCTGACAGCTGGGATGTTGGGAGGGTCC 900
Db      939 CCAGAACAGCTTGTACTGCTGATGAGATCTGACAGCTGGGATGTTGGGAGGGTCC 998
Qy      901 ATGATGATGCGGGGTGAGCTTTATATCATGGAAGCACTCTCACTATTAAAGATCTT 960
Db      999 ATGATGATGCGGGGTGAGCTTTATATCATGGAAGCACTCTCACTATTAAAGATCTT 1058
Qy      961 GGGCGGCTCCAGAGAGCTGCGGGCTGCTGAGATCGAGAGAGAGAGAGAGAGTGA 1020
Db      1059 GGGCGGCTCCAGAGAGCTGCGGGCTGCTGAGATCGAGAGAGAGAGAGAGTGA 1118
Qy      1021 GTTGTGCTTCAGATATATATGATTAACAAGTAAATATTTCCAGCTACAGTCTGGTG 1080
Db      1119 GTTGTGCTTCAGATATATATGATTAACAAGTAAATATTTCCAGCTACAGTCTGGTG 1178
Qy      1081 ATGAGTCTGACGACGAACTTTTACCCTGCGCTGCAATTCATCTGGAGTGA 1140
Db      1179 ATGAGTCTGACGACGAACTTTTACCCTGCGCTGCAATTCATCTGGAGTGA 1238
Qy      1141 GCCAGGGCTCATCATGAGAGGTTTATGAGCTGCTGAGCCCTCAATATATCATCTAGGTC 1200
Db      1239 GCCAGGGCTCATCATGAGAGGTTTATGAGCTGCTGAGCCCTCAATATATCATCTAGGTC 1298
Qy      1201 CTGAGCCATGAGAGAGACACATCAACTTTTGGATCCAGCTGAGTGGCTGAGACC 1260
Db      1299 CTGAGCCATGAGAGAGACACATCAACTTTTGGATCCAGCTGAGTGGCTGAGACC 1358
Qy      1261 AGTCTACTTATGACTTATACAGATATTTCTTCTTCATCACTCCACGAGACACCATG 1320
Db      1359 AGTCTACTTATGACTTATACAGATATTTCTTCTTCATCACTCCACGAGACACCATG 1418
Qy      1321 ACTGTCATGATCCAAAGACATGATGCTGCTGTTGGGCTGTTCTTAT 1380
Db      1419 ACTGTCATGATCCAAAGACATGATGCTGCTGTTGGGCTGTTCTTAT 1478
Qy      1381 GTTGTGACAGATGAGAAATGCTGCTAGTCC 1416
Db      1479 GTTGTGACAGATGAGAAATGCTGCTAGTCC 1514

RESULT 10
AAZ58313
ID      AAZ58313 standard, cDNA, 1884 BP.
XX
AC      AAZ58313;
XX
DT      08-MAY-2000 (first entry)
XX
DE      Human peptidase NALAD-ase IV cDNA.
XX
KW      NALAD-ase IV, N-acetylated alpha-linked acidic dipeptidase; human;
KW      chromosome 8q21.3; prostate cancer; neurodegenerative disease;
KW      Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
KW      peripheral neuropathy; Huntington's disease; acute brain injury;
KW      multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;
KW      gene therapy; diagnosis; noctropic; neuroprotective; neuroleptic;
KW      antiparkinsonian; anticonvulsant; vasotropic; ss.
XX
OS      Homo sapiens.
XX
FH      Key      location/Qualifiers
FT      CDS      149..1567
FT      /*tag= a
XX      MO200004157-A2.
XX

```

```

PD      27-JAN-2000.
XX
PE      14-JUL-1999; 99MO-GB02241.
XX
PR      14-JUL-1998; 98GB-0015284.
XX
PA      (JAN) JANSSEN PHARM NV.
XX
PI      Pangalos M, Neefs JEFM, Peeters DCG;
XX
DR      WPI; 2000-182424/16.
XX
DR      P-PSDB; AAY58879.
XX
PT      New human N-acetylated alpha-linked acidic dipeptidases for treating
PT      neural disorders e.g. Alzheimer's disease, schizophrenia and
PT      Parkinson's disease.
XX
PS      Claim 7; Fig 5; 95pp; English.
XX
CC      The present sequence is that of cDNA coding for human
CC      N-acetylated alpha-linked acidic dipeptidase IV (NALAD-ase IV,
CC      see AAY58879). The cDNA was obtained from a gall bladder cDNA
CC      library. Analysis of the open reading frame predicts a type II
CC      integral membrane protein with 5 potential N-glycosylation sites.
CC      The NALAD-ase II gene was mapped to chromosome 8q21.3.
CC      NALAD-ase IV expression was low in all tissues examined by RT-PCR.
CC      The invention provides human NALAD-ase I, II and IV polypeptides,
CC      cDNAs, antisense nucleic acids, vectors, host cells, transgenic
CC      organisms, antagonists and agonists. These are useful for treating
CC      neural disorders such as Alzheimer's disease, schizophrenia, ALS,
CC      Parkinson's disease, peripheral neuropathy, Huntington's disease,
CC      acute brain injury, multiple sclerosis, exposure to neurotoxins,
CC      peripheral nerve trauma, ischaemia or dementia (claimed). Nucleic
CC      acids can also be used for gene therapy and for genetic screening
CC      of predisposition to disorders associated with NALAD-ase.
XX
SQ      Sequence 1884 BP; 520 A; 432 C; 458 G; 474 T; 0 other;

Query Match      99.8%; Score 1412.8; DB 21; Length 1884;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ATGAAATTCCTTATCTTCGCAATTTTGGTGTGTTTACCTTTATCTGCTCTGG 60
Db      149 ATGAAATTCCTTATCTTCGCAATTTTGGTGTGTTTACCTTTATCTGCTCTGG 208
Qy      61 AAAGCTATATGACAAAGATGCACTCTTAAGAGACCTTTGAAGAAATTAAGAAATA 120
Db      209 AAAGCTATATGACAAAGATGCACTCTTAAGAGACCTTTGAAGAAATTAAGAAATA 268
Qy      121 GCCAGCTGTGAGAGATGTTGTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 180
Db      269 GCCAGCTGTGAGAGATGTTGTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 328
Qy      181 AACAGATCTATGAGCGATTTGCACTTCTGTTGATGATCTGTTGACCCAGACTGAGTGC 240
Db      329 AACAGATCTATGAGCGATTTGCACTTCTGTTGATGATCTGTTGACCCAGACTGAGTGC 388
Qy      241 TCAGAGAACTTGAAGAAAGCAATTCATTAATGATCAAAACCTGACGAGATGGCTG 300
Db      389 TCAGAGAACTTGAAGAAAGCAATTCATTAATGATCAAAACCTGACGAGATGGCTG 448
Qy      301 GAGAAATTCATCTGAGCGAGGAGATACCCCACTGGGAGAGGGAGAGAAATCAGCT 360
Db      449 GAGAAATTCATCTGAGCGAGGAGATACCCCACTGGGAGAGGGAGAGAAATCAGCT 508
Qy      361 GTGATGCTGAGGCAAGAAATTTCAATAGATAGCCATCTGGGCTTTGGACAGCATTTGG 420
Db      509 GTGATGCTGAGGCAAGAAATTTCAATAGATAGCCATCTGGGCTTTGGACAGCATTTGG 568
Qy      421 ACTCTTCAGAAAGCATTTACAGCAAGATTTCTGGTGTGAACCTTTTCATGAATCTGAG 480
Db      569 ACTCTTCAGAAAGCATTTACAGCAAGATTTCTGGTGTGAACCTTTTCATGAATCTGAG 628

```



```

QY 481 AGAAGGCTTCAGAGCAAGAGGAGAGATGTTGTTTATTAACAACCTTACATCAACTAC
DB 629 AGAAGGCTTCAGAGCAAGAGGAGAGATGTTGTTTATTAACAACCTTACATCAACTAC
QY 541 TCAAGACGGTGCATATCCAGACGAGGGGGCGGTGAGAGTCCAGAGTGGGGCTTTG
DB 689 TCAAGACGGTGCATATCCAGACGAGGGGGCGGTGAGAGTCCAGAGTGGGGCTTTG
QY 601 GCATCTCTCATTCGATCCGTGGGCTCTCTTCATCTCACTCCCTCACAAGGTATTCAG
DB 749 GCATCTCTCATTCGATCCGTGGGCTCTCTTCATCTCACTCCCTCACAAGGTATTCAG
QY 661 GAATACCAAGATGCGCTGCGCCCAAAATTCACACAGCTGTATTAAGTGAAGATGACAA
DB 809 GAATACCAAGATGCGCTGCGCCCAAAATTCACACAGCTGTATTAAGTGAAGATGACAA
QY 721 ATGATGTCAGAAATGCTTCTCATGGAGTCAAAATGTCATTCAGCTAAAGATGGGGCA
DB 869 ATGATGTCAGAAATGCTTCTCATGGAGTCAAAATGTCATTCAGCTAAAGATGGGGCA
QY 781 AAGACCTACCCAGATCTGATCTTCTCAACACTGTAGAGAGATCACTGGAGAGCAATAT
DB 929 AAGACCTACCCAGATCTGATCTTCTCAACACTGTAGAGAGATCACTGGAGAGCAATAT
QY 841 CCAGAACAGGTTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG
DB 989 CCAGAACAGGTTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG
QY 901 ATGATGATGAGCGGTGAGACCTTATATCATGAGGAGCACTCTCACTTATTAAGATCTT
DB 1049 ATGATGATGAGCGGTGAGACCTTATATCATGAGGAGCACTCTCACTTATTAAGATCTT
QY 961 GGGGTGGTCCAAAGAGAGACTCTGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
DB 1109 GGGGTGGTCCAAAGAGAGACTCTGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
QY 1021 GTTGGTGGTCCAAAGAGAGACTCTGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
DB 1169 GTTGGTGGTCCAAAGAGAGACTCTGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
QY 1081 ATGAGATCTGACGAGAGAACCTTCTTACCACTGGGCTGCAATTCATCGGAGTGAAG
DB 1229 ATGAGATCTGACGAGAGAACCTTCTTACCACTGGGCTGCAATTCATCGGAGTGAAG
QY 1141 GCCAGGGCCATCATGAGAGAGAGTATGAGCCTGCTGCAAGCCCTCAATATCACTGAGT
DB 1289 GCCAGGGCCATCATGAGAGAGAGTATGAGCCTGCTGCAAGCCCTCAATATCACTGAGT
QY 1201 CTGAGCGATGAGAGAGAGAGACATCAACTTTGAGTCCAGTGGAGTCCCTGGAGCC
DB 1349 CTGAGCGATGAGAGAGAGAGACATCAACTTTGAGTCCAGTGGAGTCCCTGGAGCC
QY 1261 AGTCTACTGATGATTAACAAGATTTCTTCTTCATCACTCCACGAGAGACCATATG
DB 1409 AGTCTACTGATGATTAACAAGATTTCTTCTTCATCACTCCACGAGAGACCATATG
QY 1321 ACTGTCATGATCCAAAGACAGATGAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCT
DB 1469 ACTGTCATGATCCAAAGACAGATGAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCT
QY 1381 GTTGTTCAGACATGAGAGAGATGCTGCTGAGTCC 1416
DB 1529 GTTGTTCAGACATGAGAGAGATGCTGCTGAGTCC 1564

```

```

RESULT 11
AA298139
ID AA298139 standard; cdna; 1923 BP.
XX AC AA298139;
XX DT 11-MAY-2000 (first entry)

```

Human signal peptide containing protein HSPB-31 cDNA seq ID NO:165.

Human, signal peptide-containing protein; HSPB, diagnosis; cancer; inflammation; cardiovascular disease; anticancer; anti-inflammatory; antimicrobial; neuroprotective; cardiovascular; hepatotropic; antiaesthetic; gene therapy; cell proliferation; neurological disorder; reproductive disorder; developmental disorder; arteriosclerosis; cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia; asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia; Parkinson's disease; Huntington's disease; ovulatory defect; muscular dystrophy; ss.

Homo sapiens.

MO200000610-A2.

06-JAN-2000.

25-JUN-1999; 99WO-0514484.

26-JUN-1998; 98US-0090762.

31-JUL-1998; 98US-0094983.

01-OCT-1998; 98US-0102686.

11-DEC-1998; 98US-0112129.

(INCY-) INCYTE PHARM INC.

Lal P, Tang YT, Gorgone GA, Corley NC, Guejler KJ, Baughn MR; Akberblom IE, Au-young J, Yue H, Patterson C, Reddy R, Hillman JU, Bandman O;

WPI: 2000-160673/14.

P-PSDS; AAY87254.

New human signal peptide-containing proteins useful in treatment, prevention and diagnosis of e.g. cancer, inflammation and cardiovascular disease.

Claim 9; Page 269-270; 327pp; English.

AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the human signal peptide-containing proteins HSPB-1 to HSPB-134. HSPBs have anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic, neuroprotective, cardiovascular and antiaesthetic activities, and can be used in gene therapy. HSPBs can be used to treat or prevent disorders associated with decreased activity or function of HSPB. Antagonists of HSPB are used to treat or prevent disorders associated with increased activity or function of HSPB. Such diseases include cell proliferation (including cancer), inflammation, cardiovascular, neurological, reproductive or developmental disorders, (e.g. arteriosclerosis, cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia, asthma, Crohn's disease, microbial or other infections, congestive or ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPB nucleic acids can be used for the recombinant production of HSPB, for detecting HSPB in standard hybridisation and amplification assays (for diagnosis and monitoring), in gene therapy, as antisense, triplex-forming or ribozyme therapeutics, for detecting related sequences or genetic variations, and for chromosomal mapping. HSPB are also used to raise specific antibodies (Ab) and to screen for agonists and antagonists (potential therapeutic agents). Ab are used to diagnose, or monitor, HSPB-related diseases (in usual immunoassays), as therapeutic antagonists, in competitive drug screens, and for purification of HSPB from natural sources.

Sequence 1923 BP; 538 A; 439 C; 471 G; 475 T; 0 other;

Query Match 99.8%; Score 1412.8; DB 21; Length 1923;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAATTCCTTACTTCGATTTTGGGTGCTTACCTTTATCCCTGCTCGGG 60


```

Db      179  ATGAAATCTTATCTTGCAGATTTTTCGATGGTGTTCACCTTTATCCCTGGCTGGG 238
Qy      61  AAAGCTATATGCAAGAAATGGCATCTCTAAGAGATTTTGAAGAAATTAAGAAATA 120
Db      239  AAAGCTATATGCAAGAAATGGCATCTCTAAGAGATTTTGAAGAAATTAAGAAATA 298
Qy      121  GCCAGCTGTGAGATGTTGTCTAAAGCAATCACTAGCTGTGTATGTAAAGCCAG 180
Db      299  GCCAGCTGTGAGATGTTGTCTAAAGCAATCACTAGCTGTGTATGTAAAGCCAG 358
Qy      181  AACAGATCTATAGACGATGGCATCTTGTGTGTATCTGTGGAACCCAGCTAGTGGC 240
Db      359  AACAGATCTATAGACGATGGCATCTTGTGTGTATCTGTGGAACCCAGCTAGTGGC 418
Qy      241  TCCAGAACTTAAAGAAAGCCATCCAAATTTATGTACCAAAACCTGACCAAGATGGCTG 300
Db      419  TCCAGAACTTAAAGAAAGCCATCCAAATTTATGTACCAAAACCTGACCAAGATGGCTG 478
Qy      301  GAGAAAGTTCACTGTGAGCCAGTGAATPACCCCACTGGAGAGAGGAGAAATCAGCT 360
Db      479  GAGAAAGTTCACTGTGAGCCAGTGAATPACCCCACTGGAGAGAGGAGAAATCAGCT 538
Qy      361  GTGATGTGAGCCAGAAATTCATTAAGATAGCCATCTGGGTCTTGGCAGCAGATTGGG 420
Db      539  GTGATGTGAGCCAGAAATTCATTAAGATAGCCATCTGGGTCTTGGCAGCAGATTGGG 598
Qy      421  ACTCTCTCAAGAAAGGATTAACAGCAGAAATTTCTGTGTGTAGCTCTTTGATGAATGAG 480
Db      599  ACTCTCTCAAGAAAGGATTAACAGCAGAAATTTCTGTGTGTAGCTCTTTGATGAATGAG 658
Qy      481  AGAAGGGCTCTGAAGAAAGAGAGAGATGTTGTTTAAACCAACCTTACATCACTAC 540
Db      659  AGAAGGGCTCTGAAGAAAGAGAGAGATGTTGTTTAAACCAACCTTACATCACTAC 718
Qy      541  TCAAGAGCGGTGCAATACCGAACGAGGCGGTGTGAAGCTGCCAGGTGGGGCTTTG 600
Db      719  TCAAGAGCGGTGCAATACCGAACGAGGCGGTGTGAAGCTGCCAGGTGGGGCTTTG 778
Qy      601  GCATCTCTCATTTCCATCCGTGGCTCTTCTTCATCTACAGTCTTCAACAGGATTCAG 660
Db      779  GCATCTCTCATTTCCATCCGTGGCTCTTCTTCATCTACAGTCTTCAACAGGATTCAG 838
Qy      661  GAATACAGAGATGGGGTCCCAAAATTCACAAGCTGTATTAAGGTGAAGATCAGAA 720
Db      839  GAATACAGAGATGGGGTCCCAAAATTCACAAGCTGTATTAAGGTGAAGATCAGAA 898
Qy      721  ATGATGTCAAGAAATGGCTTCTCATGGAGATCAAAATTTGTCATTCAGCTAAAGATGGGGCA 780
Db      899  ATGATGTCAAGAAATGGCTTCTCATGGAGATCAAAATTTGTCATTCAGCTAAAGATGGGGCA 958
Qy      781  AAGACCTTACCAAGATCTGATCTTCTCAACACTGTAGCAGAGATCTCGGAGCAATAT 840
Db      959  AAGACCTTACCAAGATCTGATCTTCTCAACACTGTAGCAGAGATCTCGGAGCAATAT 1018
Qy      841  COAGAACAGGTTGTACTGTAGTGACATCGACAGCTGGGATGTTGGGAGGGGTGCC 900
Db      1019  COAGAACAGGTTGTACTGTAGTGACATCGACAGCTGGGATGTTGGGAGGGGTGCC 1078
Qy      901  ATGATGTAGCGGTGAGACCTTTATATCATGGAGAGCACTCTCACTTATTAAGATCTT 960
Db      1079  ATGATGTAGCGGTGAGACCTTTATATCATGGAGAGCACTCTCACTTATTAAGATCTT 1138
Qy      961  GGGGTGCGTCCAAAGAGACTCTGGGCTGTGTCTGTGACATCGCAAGAAACAGAGTGA 1020
Db      1139  GGGGTGCGTCCAAAGAGACTCTGGGCTGTGTCTGTGACATCGCAAGAAACAGAGTGA 1198
Qy      1021  GTTGTGCTTCCAGATTTATAGTTACAAAGGTAATATTTCCAACTACAGTGTGAG 1080
Db      1199  GTTGTGCTTCCAGATTTATAGTTACAAAGGTAATATTTCCAACTACAGTGTGAG 1258
Qy      1081  ATGAGTCTGACGACAGAACTTCTTACCACTGGGCTGCAATTCAGTGGCAAG 1140

```

```

Db      1259  ATGAGTCTGACGACAGAACTTCTTACCACTGGGCTGCAATTCAGTGCAGTGAAG 1318
Qy      1141  GCCAGGGCATCATGAGAGAGATTATAGCTGTGACGCCCTCATATCACTCAGGTC 1200
Db      1319  GCCAGGGCATCATGAGAGAGATTATAGCTGTGACGCCCTCATATCACTCAGGTC 1378
Qy      1201  CTGAGCCATGAGAGAGAGAGACATCACTTTTGAATCCCAAGCTGAGTCCCTGAGCC 1260
Db      1379  CTGAGCCATGAGAGAGAGAGACATCACTTTTGAATCCCAAGCTGAGTCCCTGAGCC 1438
Qy      1261  AGTCTACTGATGACCTTATACAGATATTTCTTCCATCATCTCCACGAGACACATG 1320
Db      1439  AGTCTACTGATGACCTTATACAGATATTTCTTCCATCATCTCCACGAGACACATG 1498
Qy      1321  ACTGTCATGATCCCAAGCAGATGAATGTGCTGCTGCTGTTGGGCTGTTCTTAT 1380
Db      1499  ACTGTCATGATCCCAAGCAGATGAATGTGCTGCTGCTGTTGGGCTGTTCTTAT 1558
Qy      1381  GTTGTGACAGACATGAGAAATGCTGCTAGGTC 1416
Db      1559  GTTGTGACAGACATGAGAAATGCTGCTAGGTC 1594

RESULT 12
ACCS0603
ID   ACCS0603 standard; cDNA; 2077 BP.
XX
AC   ACCS0603;
XX
DT   12-JUN-2003 (first entry)
XX
DE   Human secreted protein coding sequence, SEQ ID 270.
XX
KW   Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyrostatic;
KW   vulnerary; antinflammatory; nootropic; neuroprotective;
KW   antiparkinsonian; gene therapy; human; cardiovascular disorder;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200295010-A2.
XX
PD   28-NOV-2002.
XX
PE   19-MAR-2002; 2002WO-US09785.
XX
PR   21-MAR-2001; 2001US-277340P.
PR   19-JUL-2001; 2001US-306171P.
PR   13-NOV-2001; 2001US-331287P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2003-129429/12.
XX
PT   Novel human secreted proteins, useful for detecting, preventing,
PT   diagnosing, prognosticating, treating and/or ameliorating
PT   cardiovascular disorders such as arrhythmia -
XX
PS   Claim 21; SEQ ID 270; 1881bp; English.
XX
CC   The present invention relates to novel human secreted proteins
CC   (ABR4/633-ABR49145) and their coding sequences (ACCS0344-ACCS0856). The
CC   proteins and their coding sequences are useful for the preparation of a
CC   diagnostic or pharmaceutical composition for diagnosing or treating a
CC   cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
CC   coronary arteriosclerosis and myocardial ischemia), neural disorders,
CC   immune system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, pulmonary disorders, renal disorders,
CC   proliferative disorders and/or cancerous diseases and conditions, for
CC   wound healing and epithelial cell proliferation, to treat inflammation or
CC   infection, for treating thrombosis and arteriosclerosis, for treating or

```


XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX

PI Rosen CA, Ruben SM;
 XX
 XX

DR WPI; 2003-029900/02.
 DR P-PSDB; ABR00152.
 XX

PT New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating,
 PT e.g. gastrointestinal diseases and disorders, or cancers -
 PS
 PS

XX Claim 21; Page 841-842; 1216p; English.

XX AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein-
 CC encoding cDNA clone of the invention.
 XX

SQ Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;

Query Match 99.8%; Score 1412.8; DB 25; Length 2077;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAATTCCTTATCTTGCATTTTTCGAGTGTTCACCTTTATCCCTGCTGGG 60
 DB 132 ATGAATTCCTTATCTTGCATTTTTCGAGTGTTCACCTTTATCCCTGCTGGG 191
 QY 61 AAACCTATATGCAAGAAATGSCATCTTAAGAGAACTTTGAAGAAATTAAGAAATA 120
 DB 192 AAACCTATATGCAAGAAATGSCATCTTAAGAGAACTTTGAAGAAATTAAGAAATA 251
 QY 121 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 180
 DB 252 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 311
 QY 181 AACGATCTTATGAGCGATGTCCTTCTGATGATCTTTGGAACCCAGCTAGTGGC 240
 DB 312 AACGATCTTATGAGCGATGTCCTTCTGATGATCTTTGGAACCCAGCTAGTGGC 371
 QY 241 TCCAGAACTTAAGAAAGCCATCCCAATTTATGTAACCAAACTGCGACGAAGTGGGTG 300
 DB 372 TCCAGAACTTAAGAAAGCCATCCCAATTTATGTAACCAAACTGCGACGAAGTGGGTG 431
 QY 301 GAGAAAGTTCACTGAGCCAGTGAATACCCCACTGGAGAGAGGAGAAATTCAGCT 360
 DB 432 GAGAAAGTTCACTGAGCCAGTGAATACCCCACTGGAGAGAGGAGAAATTCAGCT 491
 QY 361 GTGATGCTGAGCCAGAAATTCATTAAGATAGCAATCTCTGGGTCTTGGCAGCAGATTGGG 420
 DB 492 GTGATGCTGAGCCAGAAATTCATTAAGATAGCAATCTCTGGGTCTTGGCAGCAGATTGGG 551
 QY 421 ACTTCTCCAGAAAGCATTACGACGAAGTTCTGGTGTGACCTCTTTGATGAATCTGAG 480
 DB 552 ACTTCTCCAGAAAGCATTACGACGAAGTTCTGGTGTGACCTCTTTGATGAATCTGAG 611

QY 481 AGAAGGCTCAGAGCAAGAGGAGATGTTGTTATTAACCACTTATCATCACTAC 540
 DB 612 AGAAGGCTCAGAGCAAGAGGAGATGTTGTTATTAACCACTTATCATCACTAC 671
 QY 541 TCAAGACGGTGCATATCCGAACGCGAGGCGGTGAGCTGCCAAGTGGGGCTTTG 600
 DB 672 TCAAGACGGTGCATATCCGAACGCGAGGCGGTGAGCTGCCAAGTGGGGCTTTG 731
 QY 601 GCATCTCTCAATTTGATCCGAGGCTCCTTCTCATCTACAGTCTCCACAGGATTCAG 660
 DB 732 GCATCTCTCAATTTGATCCGAGGCTCCTTCTCATCTACAGTCTCCACAGGATTCAG 791
 QY 661 GAATACAGGATGCGTGGCCCAAAATTCACAGCTGTATTAACGTTGAGAGTGCAGAA 720
 DB 792 GAATACAGGATGCGTGGCCCAAAATTCACAGCTGTATTAACGTTGAGAGTGCAGAA 851
 QY 721 ATGATGTCAGAAATGCTTCTCATGAGGATCAAAATTTCTATTCAGTAAAGTGGGCA 780
 DB 852 ATGATGTCAGAAATGCTTCTCATGAGGATCAAAATTTCTATTCAGTAAAGTGGGCA 911
 QY 781 AAGACCTACCCAGATATCTGATTCCTTCAACATCTGACAGATCACTGGAGCAATAT 840
 DB 912 AAGACCTACCCAGATATCTGATTCCTTCAACATCTGACAGATCACTGGAGCAATAT 971
 QY 841 CCAAGACAGGTTGATCTGATGATGACATCTGACAGCTGGAGTGGGAGAGTGGC 900
 DB 972 CCAAGACAGGTTGATCTGATGATGACATCTGACAGCTGGAGTGGGAGAGTGGC 1031
 QY 901 ATGATGATGAGCGGTGAGCCTTTATATCATGAGGAGCACTCTCATTTAATGATCTT 960
 DB 1032 ATGATGATGAGCGGTGAGCCTTTATATCATGAGGAGCACTCTCATTTAATGATCTT 1091
 QY 961 GGGGTGCTCCAAAGAGGACCTGCGGCTGTGCTGTGAGCTGAGAGAAACAGTGGGA 1020
 DB 1092 GGGGTGCTCCAAAGAGGACCTGCGGCTGTGCTGTGAGCTGAGAGAAACAGTGGGA 1151
 QY 1021 GTTGTCCTTCCAGATATTAATCACTTACAGAGTAATATTTTCAACTACAGTGGTG 1080
 DB 1152 GTTGTCCTTCCAGATATTAATCACTTACAGAGTAATATTTTCAACTACAGTGGTG 1211
 QY 1081 ATGAGTCTGACGAGGAACTTTCTTACCACTGGGCTGCAATTCAGTGCAGTGAAG 1140
 DB 1212 ATGAGTCTGACGAGGAACTTTCTTACCACTGGGCTGCAATTCAGTGCAGTGAAG 1271
 QY 1141 GCCAGGCTCATATGAGAGGAGTTATGAGCTGTGAGCCCTCATATCACTCAAGTTC 1200
 DB 1272 GCCAGGCTCATATGAGAGGAGTTATGAGCTGTGAGCCCTCATATCACTCAAGTTC 1331
 QY 1201 CTGAGCCATGAGAAAGGAGCAGACATCACTTTTGTGATCCAGCTGAGTGGCAGGCC 1260
 DB 1332 CTGAGCCATGAGAAAGGAGCAGACATCACTTTTGTGATCCAGCTGAGTGGCAGGCC 1391
 QY 1261 AGTCTACTGATGATTAATCAAGATTTCTTTCATCATCTCCACGAGACACCATG 1320
 DB 1392 AGTCTACTGATGATTAATCAAGATTTCTTTCATCATCTCCACGAGACACCATG 1451
 QY 1321 ACTGTACATGATTCGAAGAGAGTAAATGTCGCTGCTGCTGTTGGGCTGTTTCTAT 1380
 DB 1452 ACTGTACATGATTCGAAGAGAGTAAATGTCGCTGCTGCTGTTGGGCTGTTTCTAT 1511
 QY 1381 GTTGTTGACAGATGAGAAATGCTGCTAGTGC 1416
 DB 1512 GTTGTTGACAGATGAGAAATGCTGCTAGTGC 1547

RESULT 14

AAK94491
 ID AAK94491 standard; cDNA; 1860 BP.

AAK94491;

06-NOV-2001 (first entry)

DE Human full-length cDNA, SEQ ID NO: 3328.
 XX
 XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX Homo sapiens.
 XX EP1130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 DR WPI: 2001-524255/58.
 DR P-PSDB; AAM93559.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesized by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a full length
 CC human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 CC
 XX
 XX
 SQ Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;
 Query Match 99.4%; Score 1408; DB 22; Length 1860;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1411; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 ATGAATTCCTTATCTTGGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 60
 DB 129 ATGAATTCCTTATCTTGGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 188
 QY 61 AAAGCTATATGCAAGATGCGATCTTGAAGGACCTTTGAAAGATTAAGAAATA 120
 DB 189 AAAGCTATATGCAAGATGCGATCTTGAAGGACCTTTGAAAGATTAAGAAATA 248
 QY 121 GCCAGCTGGAGATGTTGCTAAGCAATCACTAAGCTGTTTGTGAAGCCGAG 180
 DB 249 GCCAGCTGGAGATGTTGCTAAGCAATCACTAAGCTGTTTGTGAAGCCGAG 308
 QY 181 AACAGATCTATGAGCATTTGGCACTTCTGTTGATATCTTTGAGCCAGAGTGGC 240
 DB 309 AACAGATCTATGAGCATTTGGCACTTCTGTTGATATCTTTGAGCCAGAGTGGC 368
 QY 241 TCCAGAACTAGAAAAAGCCATCCAAATTTGTAACAAAACCTGACAGAGATGGCTG 300
 DB 369 TCCAGAACTAGAAAAAGCCATCCAAATTTGTAACAAAACCTGACAGAGATGGCTG 428
 QY 301 GAGAAATTCACCTGGAGCCAGTGAATACCCACTGGAGGAGGAGAAATTCAGCT 360
 DB 429 GAGAAATTCACCTGGAGCCAGTGAATACCCACTGGAGGAGGAGAAATTCAGCT 488
 QY 361 GTGATGCTGAGCCAGAAATTCATAGATAGCCATCTCGGTCTTGGCAGCAGCATGGG 420

DB 489 GTGATGCTGAGCCAGAAATTCATAGATAGCCATCTCGGTCTTGGCAGCAGCATGGG 548
 QY 421 ACTCCTCAGAGGCAATTAACAGAGAGTCTGTGTGACCTCTTTCAGTGAATGTCAG 480
 DB 549 ACTCCTCAGAGGCAATTAACAGAGAGTCTGTGTGACCTCTTTCAGTGAATGTCAG 608
 QY 481 AGAAGGCTCTCAGAGGCAAGAGATTTGTTTATACCAACCTTACATCACTAC 540
 DB 609 AGAAGGCTCTCAGAGGCAAGAGATTTGTTTATACCAACCTTACATCACTAC 668
 QY 541 TCAAGACGCTGCAATTCGACAGCGAGGCGGTGAGCTGCGAAGTGGGGCTTTG 600
 DB 669 TCAAGACGCTGCAATTCGACAGCGAGGCGGTGAGCTGCGAAGTGGGGCTTTG 728
 QY 601 GCATCTCTCATTTGATCCGTGGCTCTCTTCTCATCTACAGTCTCAACAGGATTCAG 660
 DB 729 GCATCTCTCATTTGATCCGTGGCTCTCTTCTCATCTACAGTCTCAACAGGATTCAG 788
 QY 661 GAATACAGAGATGGCGTCCCAAAATTCACAGCCTGTATTAACGTTGAGAAATGACAA 720
 DB 789 GAATACAGAGATGGCGTCCCAAAATTCACAGCCTGTATTAACGTTGAGAAATGACAA 848
 QY 721 ATGATGCAAGAAATGCTTCTCATGAGATCAAAATTTGTCATTCAGCTAAAGATGGGCA 780
 DB 849 ATGATGCAAGAAATGCTTCTCATGAGATCAAAATTTGTCATTCAGCTAAAGATGGGCA 908
 QY 781 AAGACCTACCCAGATCTGATTTCTTCAACATGTCAGAGATCACTGGAGCAATAT 840
 DB 909 AAGACCTACCCAGATCTGATTTCTTCAACATGTCAGAGATCACTGGAGCAATAT 968
 QY 841 CCAGAACAGTTGATCTGTCAGTGCATCTGACAGCTGGAGATGTTGGAGAGGTGCC 900
 DB 969 CCAGAACAGTTGATCTGTCAGTGCATCTGACAGCTGGAGATGTTGGAGAGGTGCC 1028
 QY 901 ATGATGATGGCGGTGAGCCTTTATATCAATGGAAGCACTTCACTTATTAAGATCTT 960
 DB 1029 ATGATGATGGCGGTGAGCCTTTATATCAATGGAAGCACTTCACTTATTAAGATCTT 1088
 QY 961 GGGGTGTCCTCAAGAGAGATCTGCGGTGCTCTGACCTGACAGAAAGAGTGA 1020
 DB 1089 GGGGTGTCCTCAAGAGAGATCTGCGGTGCTCTGACCTGACAGAAAGAGTGA 1148
 QY 1021 GTTGTGCTCTCAGATTTATAGTTACACAGGTAATTTTCAACTACAGTCTGCTG 1080
 DB 1149 GTTGTGCTCTCAGATTTATCACTTACACAGGTAATTTTCAACTACAGTCTGCTG 1208
 QY 1081 ATGAGTCTGACGAGGAACCTTCTTACCACTGGGCTGCAATTCATGCGAGTGAAG 1140
 DB 1209 ATGAGTCTGACGAGGAACCTTCTTACCACTGGGCTGCAATTCATGCGAGTGAAG 1268
 QY 1141 GCCAGGCTCATATGAGAGGTTTATGAGCTGCTGACGCCCTCAATATCATCTCAGTCTC 1200
 DB 1269 GCCAGGCTCATATGAGAGGTTTATGAGCTGCTGACGCCCTCAATATCATCTCAGTCTC 1328
 QY 1201 CTGAGCATGGAAGAGGACAGACATCACTTTTGAATCCAGGTGAGGCTGAGGCC 1260
 DB 1329 CTGAGCATGGAAGAGGACAGACATCACTTTTGAATCCAGGTGAGGCTGAGGCC 1388
 QY 1261 AGTCTACTGATGATTAACAAGTATTTCTTCCATCACTCCACGAGACACCATG 1320
 DB 1389 AGTCTACTGATGATTAACAAGTATTTCTTCCATCACTCCACGAGACACCATG 1448
 QY 1321 ACTGTCAATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1380
 DB 1449 ACTGTCAATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1508
 QY 1381 GTTGTGACAGATGAGAAATGCTGCTGAGTCC 1416
 DB 1509 GTTGTGACAGATGAGAAATGCTGCTGAGTCC 1544

RESULT 15

AAH9703
ID AAH9703 standard, cDNA, 1895 BP.
XX
AC AAH9703;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein encoding cDNA sequence SEQ ID NO:538.
XX
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiallergic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anaemia;
KW antiaggregant; haemostatic; vulnery; antilicer; osteopathic; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytotoxic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antiaphyllactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder; ss.
XX
OS Homo sapiens.
XX
PN WO20015345-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000WO-US35017.
XX
PR 23-DEC-1999; 99US-0471275.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dermanac RT;
XX
XX WPI; 2001-457603/49.
DR P-PSDB; AAM25762.
XX
XX Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS Claim 1; Page 591; 1217bp; English.
XX
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antinflammatory; antirheumatic;
CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianemic; antiaggregant; haemostatic; vulnery;
CC antilicer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytosratic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.
XX
SQ Sequence 1895 BP; 530 A; 439 C; 450 G; 476 T; 0 other;

Query Match 95.7%; Score 1355.6; DB 22; Length 1895;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1407; Conservative 0; Mismatches 9; Indels 6; Gaps 4;
QY 1 ATGAATTCCTTATCTTGGCAATTTTGGTGTGTTCACCTTTATCCTGTGCTTGGG 60
DB 126 ATGAATTCCTTATCTTGGCAATTTTGGTGTGTTCACCTTTATCCTGTGCTTGGG 185
QY 61 AAAGCTATATGCAAGATGGCATCTTAAGAGGACCTTGAAGAAATTAAGAAATA 120
DB 186 AAAGCTATATGCAAGATGGCATCTTAAGAGGACCTTGAAGAAATTAAGAAATA 245
QY 121 GCCAGCTGTGGAGATGTTGTGTAAGCAATCATCACTAGCTGTTATGTAAGCCAG 180
DB 246 GCCAGCTGTGGAGATGTTGTGTAAGCAATCATCACTAGCTGTTATGTAAGCCAG 305
QY 181 AACGATCTATGAGCGATTTGGCACTTCTGTGTGATACCTGTGAGCCCAAGCTGAGTGGC 240
DB 306 AACGATCTATGAGCGATTTGGCACTTCTGTGTGATACCTGTGAGCCCAAGCTGAGTGGC 365
QY 241 TCCAGAAACCTAAGAAAAGCCATCCAAATTAATGACCAAAACCTGCAGCAAGAGGGCTG 300
DB 366 TCCAGAAACCTAAGAAAAGCCATCCAAATTAATGACCAAAACCTGCAGCAAGAGGGCTG 425
QY 301 GAGAAAGTTCACTGAGCCAGTGAATACCCCACTGGAGAGGGAGAAATCAAGCT 360
DB 426 GAGAAAGTTCACTGAGCCAGTGAATACCCCACTGGAGAGGGAGAAATCAAGCT 485
QY 361 GTGATGTGAGACCAAGAAATTCATAGATAGCCATCTGTGGTCTTGGCAGCAGCTTGGG 420
DB 486 GTGATGTGAGACCAAGAAATTCATAGATAGCCATCTGTGGTCTTGGCAGCAGCTTGGG 545
QY 421 ACTGCTCCAGAAAGCATTAAGCAGAAAGTCTGTGTGACCTTTTCATGAACTGAC 480
DB 546 ACTGCTCCAGAAAGCATTAAGCAGAAAGTCTGTGTGACCTTTTCATGAACTGAC 605
QY 481 AGAAGGCTCAGAAAGCAAGAGGAAATGTTGTTATATACCAACCTTACATCACTAC 540
DB 606 AGAAGGCTCAGAAAGCAAGAGGAAATGTTGTTATATACCAACCTTACATCACTAC 665
QY 541 TCAAGACGGTGCATACCGAAGCGCAGGGGCGGTGAGCTGCAGAGTGGGGCTTTTG 600
DB 666 TCAAGACGGTGCATACCGAAGCGCAGGGGCGGTGAGCTGCAGAGTGGGGCTTTTG 725
QY 601 GCATCTCTCATTTGATTCGCTGGGCTCTCTTCCATCTACAGTCTCTCAACAGATTCAG 660
DB 726 GCATCTCTCATTTGATTCGCTGGGCTCTCTTCCATCTACAGTCTCTCAACAGATTCAG 785
QY 661 GAATACCAAGATGGCGGCCCAAAATTCGAACAGCCTGTATACGGTGAAGATGCAGAA 720
DB 786 GAATACCAAGATGGCGGCCCAAGATTCGAACAGCCTGTATACGGTGAAGATGCAGAA 845
QY 721 ATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGCA 780
DB 846 ATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGCA 905
QY 781 AAGACCTACCAAGATGATGATTTCTTCAACATCTGTAGCAAGATCATCTGGAGCAAAATAT 840
DB 906 AAGACCTACCAAGATGATGATTTCTTCAACATCTGTAGCAAGATCATCTGGAGCAAAATAT 965
QY 841 CCAAGACAGGTTGATCTGTCAGTGCATCTGGAACAGCTGGAGTGTGGGAGGGTGGC 900
DB 966 CCAAGACAGGTTGATCTGTCAGTGCATCTGGAACAGCTGGAGTGTGGGAGGGTGGC 1025
QY 901 ATGATGATGCGGTGAGCGCTTTATATCATGGAAGACCTTCATTTATTAAGATCTT 960
DB 1026 ATGATGATGCGGTGAGCGCTTTATATCATGGAAGACCTTCATTTATTAAGATCTT 1085
QY 961 GGGCTGGCTCCAAAGAGAACTTCGGGCTGTGCTGACCTGCAAGAAACAAGTGGGA 1020
DB 1086 GGGCTGGCTCCAAAGAGAACTTCGGGCTGTGCTGACCTGCAAGAAACAAGTGGGA 1145
QY 1021 GTTGTGCTCTTCCAGATTTATTCAGTTTACCAAGGTAAATTTTCCAACTACGTGTG 1080

```
Db      ||||| 1146 GTTGTGCTTCCAGTATTACAGTTACACAAAGGTAATATTCCAACTACAGTCTGCTG 1205
QY      ||||| 1081 ATGAGTCTGACGAGGAACTTCTTACCCACTGGGCTGCAATTCACTGGCAGTGAAGAAG 1140
Db      ||||| 1206 ATGAGTCTGACGAGGAACTTCTTACCCACTGGGCTGCAATTCACTGGCAGTGAAGAAG 1265
QY      ||||| 1141 GCCAGGCGCATCATGAGAGGTTATGAGCTGCTGACAGCCCTCAATATCATCTCAGGTC 1200
Db      ||||| 1266 GCCAGGCGCATCATGAGAGGTTATGAGCTGCTGACAGCCCTCAATATCATCTCAGGTC 1325
QY      ||||| 1201 CTGAGCCATGAGAGAGGACACATCACTTTTGATCCAAGCTGAGTGCTTGAGGCC 1260
Db      ||||| 1326 CTGAGCCATGAGAGAGGACACATCACTTTTGATCCAAGCTGAGTGCTTGAGGCC 1385
QY      ||||| 1261 AGTCTACTGATGACTTATAGAAATTTCTTCCATCACTCCACGAGACACCATG 1320
Db      ||||| 1386 AGTCTACTGATGACTTATAGAAATTTCTTCCATCACTCCACGAGACACCATG 1445
QY      ||||| 1321 ACTGTC--ATGATCCAAAGCAGATGATG-TTGCTGCTGCTG-TTGAGGCTGTGTTT 1375
Db      ||||| 1446 ACTGTCATGGGATCCAAACGCAATGATGTTGCTGCTGCTGTTTGGGCTGTGTTT 1505
QY      ||||| 1376 CTATATGT-TGTTGACAGACATGAAAGAAATGCTGCTAGGTCC 1416
Db      ||||| 1506 CTATATGTGTTGACAGACATGAAAGAAATGCTGCTAGGTCC 1547
```

Search completed: December 22, 2003, 15:10:01
Job time : 346.284 secs

THIS PAGE BLANK (USPTO)

Tue Dec 23 11:24:35 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:22:34 ; Search time 2771.56 Seconds
(without alignments)
12417.225 Million cell updates/sec

Title: US-09-745-763-35_COPY_99_1514
Perfect score: 1416
Sequence: 1 ATGAAATTCCTTACTTCGC.....AGGAATGCTGCTAGSTCC 1416

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1118.8	79.0	1868	11	BC010977 Mus muscu
2	1116.8	78.9	1739	11	AK032972 Mus muscu
3	1115.2	78.8	1805	11	AK075686 Mus muscu
4	958.8	67.7	1024	11	BC012019 Homo sapi

5	905.8	64.0	1050	13	BX439467
6	843.8	59.6	1088	13	BX355940
7	796.8	56.3	909	13	BQ878966
8	774.4	54.7	1126	13	BX360507
9	766.2	54.1	919	13	BX335995
10	748.6	52.9	866	13	BX146905
11	742.8	52.5	1012	13	BX416896
12	740	52.3	902	10	BG751497
13	736.6	52.0	998	13	BX460463
14	726.6	51.3	996	13	BX355939
15	719.2	50.8	739	13	CB958693
16	683.6	48.3	813	14	CD519131
17	683.4	48.3	922	12	BG761741
18	681.8	48.1	708	14	CD366462
19	676.6	47.8	945	12	BI909780
20	666.8	47.1	1081	13	BQ072892
21	665.2	47.0	889	10	BE906771
22	663.6	46.9	801	12	BI754143
23	655.8	46.3	701	12	BI771109
24	653.2	46.1	1020	13	BX439466
25	642.4	45.4	786	10	BG430966
26	641.8	45.3	783	12	BI754468
27	636.8	45.0	653	10	BE218907
28	632.8	44.7	822	12	BI911772
29	630.4	44.5	650	10	BG432644
30	627.8	44.3	3151	11	BC017373
31	626.2	44.2	804	14	CD352906
32	626.2	44.2	879	13	BQ218838
33	625.6	44.2	677	10	BE670671
34	620	43.8	944	12	BI819273
35	614.8	43.4	1073	13	BX360506
36	604	42.7	926	13	BQ938234
37	599.6	42.3	805	10	BG563740
38	592.6	41.9	1201	13	BX379674
39	591.6	41.8	849	10	BG166321
40	591.2	41.8	1003	13	BX416895
41	590.8	41.7	751	14	CD558931
42	590.8	41.7	806	12	BI764052
43	590.2	41.7	727	12	BG765433
44	583.8	41.2	736	12	BI769575
45	577.8	40.8	745	12	BI462584

ALIGNMENTS

RESULT 1
LOCUS BC010977
DEFINITION Mus musculus, similar to plasma glutamate carboxypeptidase, clone IMAGE:4018296, mRNA.
ACCESSION BC010977
VERSION BC010977.1
KEYWORDS GI:15012137
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1868)
AUTHORS Strausberg, R.
JOURNAL Direct Submission
Submitted (23-UTL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, O., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadori, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schirml, L. M., Steubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamliya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mommaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kontseki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1739)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome

FEATURES
SOURCE

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
location/Qualifiers

1. 1739
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:6720483C06"
/db_xref="taxon:10090"
/clone="6720483C06"
/sex="male"
/tissue_type="wolfian duct includes surrounding region"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
81. 1493
/note="unnamed protein product; plasma glutamate carboxypeptidase (MGI|MGI:1889205, GB|NM_018755, evidence: BLASTN, 99%, match=943)
putative"
/codon_start=1
/protein_id="BAC28105.1"
/db_xref="GI:2632873"
/translation="MSLFLFLVHLAAGCAVRKNGVSGTFRFEIKELNEDV AKAIINLAVYGRKONSREYELIVITVGRSLGSKNLEAIDIMVONQDLENVH LEOVRIPHMERGESAVMLEPRIRHNAIIGLSSIGIPPGIATVIVVSPDELQR ASARGKIIVYNOPTGYEKTQYRYOGAAVKAAGVAVSLQSVASFISYHTGIO KYDGGPKIPFATITVEDAMSRMNRNGKIVIHLEMAKTYPDSDPNTVAETGS MYREVVIVUSQTEIDSWNDGALDDGGAATSEALSIYVDLRLPRTRILYMTAE EGGTGAQSYTEIHRANISRTSLWEDSFTFPTGQFTGSDKAPRIMEVNLQAP LNTGTFNSGEGDINFWIQAGVGLSRDLVLYKFFHSHGDTWMDPKOMVVA AVAAVVAIVVADMDLPR8"
1722. 1727
/note="putative"
1739
/note="putative"
polya_signal
polya_site
BASE COUNT 482 a 369 c 423 g 465 t
ORIGIN

Query Match 78.9%; Score 1116.8; DB 11; Length 1739;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1217; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

33 TGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATGCAAGATGCATCTCTAAGAG 92
107 TGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATGCAAGATGCATCTCTAAGAG 166
93 GACTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGAGATGTTCCTAAAGCATCAT 152
167 AACCTTTCGAGAAATTAAGAAAGAAATAGCCAGCTGTGAGATGTTCCTAAAGCATCAT 226
153 CAACCTGCTTTTATGTTAAGCCGAGAACAGATCTTAAGAGCATTTGCACTTCTGCT 212
227 CAACCTGCTTTTATGTTAAGCCGAGAACAGATCTTAAGAGCATTTGCACTTCTGCT 286
213 TGAATGTTGAGCCCGAGCTGAGTGGCTCCAGAAACCTTAAGAAACCTCAATATAT 272
287 TGAATGTTGAGCCCGAGCTGAGTGGCTCCAGAAACCTTAAGAAACCTCAATATAT 346
273 GTACCAAAACCTGAGCAAGATGGCTGAGAAAGTTCACTTGAAGCCAGTGAAGATAC 332
347 GTACCAAAACCTGAGCAAGATGGCTGAGAAAGTTCACTTGAAGCCAGTGAAGATAC 406
333 CCACCTGGAGAGGGGAGAAAGATCAGCTGTGATCTGTGAGCCAGAAATTTGTAAGATAC 392
407 CCACCTGGAGAGGGGAGAAAGATCAGCTGTGATCTGTGAGCCAGAAATTTGTAAGATAC 466
393 CATCTGGCTTGTGAGCAGCATTTGGGATCTCTCCAGAGAGCATTTACAGAGATGTTCT 452

```

Db 467 TATTCTAGTCTTGGCAGCAGCATTTGGAGCTCCCCAGAGGACATCACAGAGAAGTCT 526
Oy 453 GGTGTGACCTCTTTCGATGAATCGACAGAGAGGCTCCAGAAAGAAAGGAAGATGT 512
Db 527 GGTGTGACCTCTTTCGATGAATCGACAGAGAGGCTCCAGAAAGAAAGGAAGATGT 586
Oy 513 TGTATTAACCAACCTTATCATCACTACTCAAGAGCGGTGCATACCAAGCAGAGGAGC 572
Db 587 TGTATTAACCAAGCTTATCATCTGCTATAGAGAAAGCTGTGAGTACCGGGTGCAGAGC 646
Oy 573 GGTGAGACCTCCCAAGGTGGGGGCTTTGGCATCTCTCATTTGCATCGTCCCTCTTC 632
Db 647 TGTGGAAGCTCCCAAGGTGGAGGCTGTGGCATCTCTCATCAATCAGTAGCTCTTTC 706
Oy 633 CATCTACAGTCTCTCACACAGGATTCAGAGATACAGAGATGGCCGCCAAAATTCAC 692
Db 707 CATCTACAGTCTCTCACACAGGATTCAGAGATACAGAGATGGCCGCCAAAATTCAC 766
Oy 693 AGCCTGATTTACGGTGAAGATGACAGAAATGATGTCAAGAAATGGCTCTCATGGATCA 752
Db 767 AGCCTGATTTACAGTGAAGATGACAGAAATGATGTCTGAATGGCTCTCTGGGAAACA 826
Oy 753 AATTGTCAATCACTGTAAGATGGGGGCAAGACCTACCCAGATCTGATCTCTTCAAC 812
Db 827 AATTGTCAATCACTGTAAGATGGGGGCAAGACCTACCCAGATCTGATCTCTTCAATAC 886
Oy 813 TGTACAGAGATCACTGGGAGAGATATTCACAGAGATGATGTCACTGTCTGATGAGCATCT 872
Db 887 TGTGTGACAGATCACTGGGAGAGATATTCACAGAGATGATGTCTGTGTGAGGACATTT 946
Oy 873 GACAGAGTGGAGTGTGGGAGAGGAGTGCATGATGATGAGGAGCTTATATCATG 932
Db 947 GACAGAGTGGAGTGTGGGAGAGGAGTGCATGATGATGAGGAGCTTATATCATG 1006
Oy 933 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGAGACTCTGGGCTGT 992
Db 1007 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGAGACTCTGGGCTGT 1066
Oy 993 GCTCTGACCTGACAGAGAAAGAGTGGAGTGTGTCTTCCAGTATTCAGTACACA 1052
Db 1067 GCTCTGACCTGACAGAGAAAGAGTGGAGTGTGTCTTCCAGTATTCAGTACACA 1126
Oy 1053 GGTAAATATTTTCAACTACAGTCTGTGATGAGTGTGACGAGAGACCTTCTTACCCAC 1112
Db 1127 GGTAAATATTTTCAACTACAGTCTGTGATGAGTGTGACGAGAGACCTTCTTACCCAC 1186
Oy 1113 TGGGCTGCAATTCACCTGACAGTGAAGAGCCAGGCGCATCATGAGAGAGTTATGAGCTT 1172
Db 1187 TGGACTGCAAGTTCACTGGCAGTGAAGAGCCAGGCGCATCATGAGAGAGTTATGATCT 1246
Oy 1173 GGTGAGAGCCCTCATATTCATCTCAGTCTGTAGCCATGAGAGAGGACAGACATCACTT 1232
Db 1247 TGTGCAACCCCTCATATTCATCTCAGTCTGTAGCCATGAGAGAGGACATGATTAATCTT 1306
Oy 1233 TTTGATCAAGCTGAGTGTGCTGGAGCCAGTCTACTTGTATGACTATTAACAATTTCTT 1292
Db 1307 CTGATTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1366
Oy 1293 CTTCATCACTCTCCACGAGAGACCATGATCTGTCAATGATCCAAAGAGATGATGTC 1352
Db 1367 TTTTCATCACTCTCCACGAGAGACCATGATCTGTCAATGATCCAAAGAGATGATGTC 1426
Oy 1353 TGTCTCTGTTGGGCTGTTGTTTCTTATGTTGTTCAGACATGAGAAAGTCTCTCAG 1412
Db 1427 TGTCTCTGTTATGGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1486
Oy 1413 GTCC 1416
Db 1487 GTCC 1490

```

RESULT 3
AK075686

LOCUS AK075686 1805 bp mRNA linear HTC 07-DEC-2002
DEFINITION Mus musculus 18-day embryo whole body cDNA, RIKEN full-length
cDNA library, clone:1190003p2 product:plasma glutamate
carboxypeptidase, full insert sequence.

ACCESSION AK075686
VERSION AK075686.1 GI:26344480
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORANISM Mus musculus

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20493374
PUBMED 11042159

REFERENCE
AUTHORS
1 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Komno, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS
4 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., Kiro, C., Kouchiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pezole, G., Quackenbush, J.,
Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T.,
Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barah, G.,
Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F.,
Brownstein, M. J., Bult, C., Fletcher, P., Fujita, M., Gariboldi, M.,
Gustincich, S., Hill, D., Holtmann, M., Hume, D. A., Kamiya, M.,
Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, D., Mombaerts, P., Nordone, P.,
Rising, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K.,
Schonbach, C., Seta, T., Shibata, Y., Schorch, K. F., Suzuki, H., Toyo-oka, K.,
Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A.,
Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S., and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851

REFERENCE
AUTHORS
5 The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 1217851
PUBMED 1217851

REFERENCE
AUTHORS
6 (bases 1 to 1805)
Aizawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,

Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Yoshida, Y., Toyota, T., Yamamura, T., Yamane, I., Yanunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, FAX: 81-45-503-9216)

COMMENT

Genomic library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES
source

1. 1805
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:1190003P12"
/db_xref="MGI:1909852"
/db_xref="taxon:10090"
/clone="1190003P12"
/issue_type="whole body"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18-day embryo"
147..1559
/note="unnamed protein product; plasma glutamate carboxypeptidase (MGP)|MGI:1889205, GB|NM_018755, evidence: BLASTN, 99%, match=943)
putative"
/codon_start=1
/protein_id="BAC35891.1"
/db_xref="GI:26344481"
/db_xref="MGI:1923585"
/transcript="MRSFLPFLVHLALSGKAVPKXVSORTFREIKBIANYEDV AKAIINLAIVGKYNRSYERGLVLDVGPRLSGSKNLKRAI QMYONLQDDLENVH LEQVRIPMERGESAVMLEPR IHKMA ILGLSSIGTPPGITAEVIVASFDLQRL ASEARGLIYVNOPTGVEKTVQYVOGAFAAKVAVASLISVASFSIYSPHTIO KYODGVKIPACTIVDAEMMSMSMRKNTIVHLMEGAKTVPDTSFMTVAITGS MYREVVIVSGHLDSPVGGALDDGGGATISWELSLVYKDLGRPKRTLEVLMTAE EQGIGASQYIELHKANISKYSLVWEDSGFLPTGLQFTGSDVARIMKEVNLLOP LNVTKVSNSEGTIDINFMIOAGVPGASLRDLDLYKFFHHSHGDTMTVMDPKONVAA AVMAVVAIVVADMDMLPRS"
1788..1793
/note="putative"
1805
/note="putative"

CDS

polya_signal
1805
/note="putative"
polya_site
1805
/note="putative"
BASE COUNT 486 a 397 c 445 g 476 t 1 others
ORIGIN

Query Match 78.8%; Score 1115.2; DB 11; Length 1805;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1216; Conservative 0; Mismatches 166; Indels 0; Gaps 0;
QY 33 TGTTCACCTTTATCCCTGCTGCTGGGAAGCTATATGCAAGATGCACTCTTAAGAG 92
Db 173 TGTTCACCTTTATCCCTGCTGCTGGGAAGCTATATGCAAGATGCTTTTCTCAGCG 232
QY 93 GACTTTGAAGAAATTAAGAAATGCAAGCTGTGGAGATGTTGCTAAAGCAATCAT 152
Db 233 AACATTCGAGAAATTAAGAAATGCAAGCTATGATGATGTTGCTAAAGCAATCAT 292

QY 153 CAACCTAGCTGTTATGTTAAGCCCAAGAACAGATCTATAGCGATGGCACTTCTGCT 212
Db 293 CAACCTAGCTGTTATGTTAAGAACCAAGAACCGGCTCTATAGCGATGGCACTTCTGCT 352
QY 213 TGATACCTGTTGACCCAGACTGATGCTCTCAAGAACTTGAAGAAAGCCATCCAAATAT 272
Db 353 TGATACCTGTTGACCCAGACTGATGCTCTCAAGAACTTGAAGAAAGCCATCCAAATAT 412
QY 273 GTACCAAAACCTGACAGCAATGAGGCTGGAGAAAGTTCACTGGAGCAGTGAATATAC 332
Db 413 GTACCAAAACCTGACAGCAATGAGGCTGGAGAAAGTTCACTGGAGCAGTGAATATAC 472
QY 333 CCACTGGAGAGAGGAGAAAGCAATGAGTGTGATGCTGAGCCCAAGATTTATAGATAGC 392
Db 473 CCACTGGAGAGAGGAGAAAGCAATGAGTGTGATGCTGAGCCCTGAGATTTCAAGATAGC 532
QY 393 CATCTGGGCTTTGGCAGCAGCAATTTGGGACTCTCCAGAAAGCATTAACAGAGATTC 452
Db 533 TATTTGAGTCTTTGGCAGCAGCAATTTGGGACTCTCCAGAGAGCATCAACAGATGCT 592
QY 453 GGTGTGAGCTCTTTGGAGATGAGTGAAGAGGAGGCTCAGAAAGCAAGAGGAGATGT 512
Db 593 GGTGTGAGCTCTTTGGAGATGAGTGAAGAGGAGGCTCAGAAAGCAAGAGGAGATGT 652
QY 513 TGTATTAAACCAACTTACATCACTCAAGGACGCTCAATACCGAAGCGAGGAGG 572
Db 653 TGTATTAAACCAAGCTTACATCACTGAGTGAAGAGGAGGCTCAGAAAGCGAGGAGG 712
QY 573 GGTGTGAGCTCTTTGGAGATGAGTGAAGAGGAGGCTCAGAAAGCGAGGAGG 632
Db 713 TGTGTGAGCTCTTTGGAGATGAGTGAAGAGGAGGCTCAGAAAGCGAGGAGG 772
QY 633 CATCTAGGCTCTCAACAGATTTCAAGAAATTCAGAAAGGAGGAGGAGGAGGAGG 692
Db 773 CATCTAGGCTCTCAACAGATTTCAAGAAATTCAGAAAGGAGGAGGAGGAGGAGG 832
QY 693 AGCTGTATTACGCTGAGGAGATGAGTCAAGAAATGATGATGATGATGATGATGAT 752
Db 833 AGCTGTATTACGCTGAGGAGATGAGTCAAGAAATGATGATGATGATGATGATGAT 892
QY 753 AATTGTCAATCACTGAGAAATGAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 812
Db 893 AATTGTCAATCACTGAGAAATGAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 952
QY 813 TGTAGCAGATCACTGAGGAGCAATTCAGAAAGAGTGTGATGATGATGATGATGAT 872
Db 953 TGTAGCAGATCACTGAGGAGCAATTCAGAAAGAGTGTGATGATGATGATGATGAT 1012
QY 873 GGAAGCTGGAGATGTTGGGAGGAGGCTCATGATGATGATGATGATGATGATGATG 932
Db 1013 GGAAGCTGGAGATGTTGGGAGGAGGCTCATGATGATGATGATGATGATGATGATG 1072
QY 933 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCAAGAGAGCTCTGGGCTGCT 992
Db 1073 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCAAGAGAGCTCTGGGCTGCT 1132
QY 993 GCTTGTGACCTGACAGAAACAAGGTGAGTGTGCTTCCAGTATTTATCACTTAC 1052
Db 1133 GCTTGTGACCTGACAGAAACAAGGTGAGTGTGCTTCCAGTATTTATCACTTAC 1192
QY 1053 GGTAAATATTTTCAACACTGAGTGTGATGATGATGATGATGATGATGATGATGAT 1112
Db 1193 GGTAAATATTTTCAACACTGAGTGTGATGATGATGATGATGATGATGATGATGAT 1252
QY 1113 TGGGCTCAATTCACCTGAGAGTGAAGAGGCAAGGAGGCTCATGAGAGAGGTTATG 1172
Db 1253 TGGGCTCAATTCACCTGAGAGTGAAGAGGCAAGGAGGCTCATGAGAGAGGTTATG 1312
QY 1173 GCTGACAGCCCTCAATTCACCTGAGTGTGATGATGATGATGATGATGATGATGAT 1232
Db 1313 TCTGCAACCCCTCAATTCACCTGAGTGTGATGATGATGATGATGATGATGATGAT 1372
QY 1233 TTGATCAAGCTGAGTGTGAGGAGGCTGATGATGATGATGATGATGATGATGATG 1292

```

Db      1373  CTGATCCAGCTGAGTACCGAGCCAGCTGCTGATGATCTGTCAAGTATTTCTT 1432
Oy      1293  CTTCCATACCTCCCAACGAGACACATGATCTGTATGATCCAAAGCATGATGATTTGC 1352
Db      1433  TTTCCATACCTCCCAACGAGACACATGATCTGTATGATCCAAAGCATGATGATTTGC 1492
Oy      1353  TGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGGAAGAAATGCTGCTAG 1412
Db      1493  TGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGGAAGAAATGCTGCTAG 1552
Oy      1413  GTCC 1416
Db      1553  GTCC 1556

```

```

RESULT 4
BC012019      1024 bp mRNA linear HTC 04-MAR-2003
LOCUS      Homo sapiens, similar to plasma glutamate carboxypeptidase, clone
DEFINITION IMAGE:4455631, mRNA.
ACCESSION BC012019.1 GI:15080558
VERSION BC012019.1
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1024)
AUTHORS Strausberg, R.
TITLE Direct Submision
JOURNAL Submitted (30-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK NIH-MGC Project URL: http://mgc.mci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnarathne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Louised, H.,
Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
 Series: IRAX Plate: 28 Row: b Column: 11
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5174626
 This clone has the following problem: retained intron.

FEATURES

```

SOURCE
1..1024
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4455631"
/cisue_type="kidney, hypernephroma"
/clone_id="NIH_MGC_89"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
BASE COUNT 317 a 210 c 256 g 241 t
ORIGIN

```

```

Query Match 67.7%; Score 958.8; DB 11; Length 1024;
Best Local Similarity 99.8%; Pred. No. 1.4e-269;
Matches 960; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy      1 ATGAATTCCTTATCTTGCATTTTGGTGTGTTGACCTTTATCTGCTGTGG 60
Db      36 ATGAATTCCTTATCTTGCATTTTGGTGTGTTGACCTTTATCTGCTGTGG 95
Oy      61 AAAGCTATATGCAAGAAATGGCATCTTAAGAGACCTTTGAGAAATAAAGAAATA 120
Db      96 AAAGCTATATGCAAGAAATGGCATCTTAAGAGACCTTTGAGAAATAAAGAAATA 155
Oy      121 GCCAGCTGTGAGATGTGTGCTTAAAGCAATCAACCTAGCTGTTTATGTTAAAGCCAG 180
Db      156 GCCAGCTGTGAGATGTGTGCTTAAAGCAATCAACCTAGCTGTTTATGTTAAAGCCAG 215
Oy      181 AACAGATCTATAGACGATTTGGCATTCTGTTGATCTGTGAGACCACTGAGTGGC 240
Db      216 AACAGATCTATAGACGATTTGGCATTCTGTTGATCTGTGAGACCACTGAGTGGC 275
Oy      241 TCCAAGAACTTGAAGAAAGCCATCCAAATTAATGTAACCAAACTGACAGAAATGGCTG 300
Db      276 TCCAAGAACTTGAAGAAAGCCATCCAAATTAATGTAACCAAACTGACAGAAATGGCTG 335
Oy      301 GAGAAATTTACCTTGAGCCAGTGAAGTAACCCCACTGGGAGAGGGAGAAATCAAGT 360
Db      336 GAGAAATTTACCTTGAGCCAGTGAAGTAACCCCACTGGGAGAGGGAGAAATCAAGT 395
Oy      361 GTGATGCTGAGCCAGAAATTCATAGATAGCCATCTGGGCTTGGCAGCAGATTTGG 420
Db      396 GTGATGCTGAGCCAGAAATTCATAGATAGCCATCTGGGCTTGGCAGCAGATTTGG 455
Oy      421 ACTCTCCAGAAAGCATTTACAGCAAGAGTTCTGTGTGTAACCTTTTCATGAACTGAG 480
Db      456 ACTCTCCAGAAAGCATTTACAGCAAGAGTTCTGTGTGTAACCTTTTCATGAACTGAG 515
Oy      481 AGAAGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      516 AGAAGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 575
Oy      541 TCAAGACGCTGCAATACCAAGCAGAGGCGGTGAGAGCTGCCAAGTGGGGCTTTG 600
Db      576 TCAAGACGCTGCAATACCAAGCAGAGGCGGTGAGAGCTGCCAAGTGGGGCTTTG 635
Oy      601 GCATCTCTCATTTGATCGTGGCTCTCTTCTCATCTACAGTCTCAACAGTATTGAG 660
Db      636 GCATCTCTCATTTGATCGTGGCTCTCTTCTCATCTACAGTCTCAACAGTATTGAG 695
Oy      661 GAATACAGAGTGGCGGCCCAATTCGAAGCCGTATTACGTGGAGAGAGAGAGAG 720
Db      696 GAATACAGAGTGGCGGCCCAATTCGAAGCCGTATTACGTGGAGAGAGAGAGAG 755
Oy      721 ATGATGTCAAGAAATGGCTTCTCATGAGATCAAAATTTGATTCAGCTTAAAGATGGGCA 780
Db      756 ATGATGTCAAGAAATGGCTTCTCATGAGATCAAAATTTGATTCAGCTTAAAGATGGGCA 815
Oy      781 AAAGCTACCCAGATAGTATTCCTTCAACACTGTAGACAGATCACTGGAGCAATAT 840
Db      816 AAAGCTACCCAGATAGTATTCCTTCAACACTGTAGACAGATCACTGGAGCAATAT 875
Oy      841 CCAAGACAGTTGTACTGTGTCAGTGAACATCTGACAGCTGGAGATGTTGGCAGGGTGC 900
Db      876 CCAAGACAGTTGTACTGTGTCAGTGAACATCTGACAGCTGGAGATGTTGGCAGGGTGC 935
Oy      901 ATGATGATGGCGGTGAGCCTTATATATCAATGGAAGACATCTCACTTATTAAGATCTT 960
Db      936 ATGATGATGGCGGTGAGCCTTATATATCAATGGAAGACATCTCACTTATTAAGATCTT 995
Oy      961 GG 962
Db      996 GG 997

```

```

RESULT 5
BX439467      1050 bp mRNA linear EST 15-MAY-2003
LOCUS      Homo sapiens PLACENTA Homo sapiens cDNA clone CSDB010YJ06
DEFINITION BX439467

```


5-PRIME, mRNA sequence.
BX439467
EST.
BX439467.1 GI:30777746
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1050)
L1, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE010DE03QPI&cluster=4663.f. Contact :
Feng liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE010DE03QPI.
Location/Qualifiers
1. 1050
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE010Y006"
/tissue_type="PLACENTA"
/clone_id="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 307 a 233 c 271 g 239 t
ORIGIN
Query Match 64.0%; Score 905.8; DB 13; Length 1050;
Best Local Similarity 99.8%; Pred. No. 5.2e-254;
Matches 907; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAATTCCTTATCGATTTTGGGTGCTTACCTTTATCCCTGCTGGG 60
DB 142 ATGAATTCCTTATCGATTTTGGGTGCTTACCTTTATCCCTGCTGGG 201
QY 61 AAAGCTATGCAAGATGCGATCTTAAAGAGACTTTTGAAGAAATTAAGAATA 120
DB 202 AAAGCTATGCAAGATGCGATCTTAAAGAGACTTTTGAAGAAATTAAGAATA 261
QY 121 GCCAGCTGTGAGATGTTGCTAAAGCAATCAACCTAGCTTTTATGTTAAAGCCAG 180
DB 262 GCCAGCTGTGAGATGTTGCTAAAGCAATCAACCTAGCTTTTATGTTAAAGCCAG 321
QY 181 AAGCATCCATGAGCGATGGGCACTCTGTTGATATCTGTTGAACCAAGCTAGTGGC 240
DB 322 AAGCATCCATGAGCGATGGGCACTCTGTTGATATCTGTTGAACCAAGCTAGTGGC 381
QY 241 TCCAAAGACTTAAGAAAAGCATCAATTAATACCAAACTGACAGCAATGGGCTG 300
DB 382 TCCAAAGACTTAAGAAAAGCATCAATTAATACCAAACTGACAGCAATGGGCTG 441
QY 301 GAGAAAGTTCAGTGAAGCGAGTGAAGTACCCCACTGGAGAGAGGAGAAATCAAGT 360
DB 442 GAGAAAGTTCAGTGAAGCGAGTGAAGTACCCCACTGGAGAGAGGAGAAATCAAGT 501
QY 361 GTGATGCTGAGCGCAAAATTCATTAAGTACCAATCCCTGGGCTTTGGAGAGCAATGGG 420
DB 502 GTGATGCTGAGCGCAAAATTCATTAAGTACCAATCCCTGGGCTTTGGAGAGCAATGGG 561
QY 421 ACTCTCCAGAAAGCATTAAGAGAGTTCGTGATGATCCTTTTCATGATGAGTGGAG 480

|||||
DB 562 ACTCTCCAGAAAGCATTAAGAGAGTTCGTGATGATCCTTTTCATGATGAGTGGAG 621
QY 481 AGAAGGGCTTCAGAAAGAGAGAAATGTTGTTTATTAACCACTTTCATCACTAC 540
DB 622 AGAAGGGCTTCAGAAAGAGAGAAATGTTGTTTATTAACCACTTTCATCACTAC 681
QY 541 TCAAGAGCGTGCATATCCAGACGAGGGGCGGTGAGAGTGCAGAGTGGGGCTTTG 600
DB 682 TCAAGAGCGTGCATATCCAGACGAGGGGCGGTGAGAGTGCAGAGTGGGGCTTTG 741
QY 601 GCATCTCATTTGATCGATCGGCTCTCTTTCATCACTAGTCTCAACAGTATTAG 660
DB 742 GCATCTCATTTGATCGATCGGCTCTCTTTCATCACTAGTCTCAACAGTATTAG 801
QY 661 GAATACAGAGATGGCGGCCCAAAATTCACAGCCCTGTTATAGTGGAGAGTGCAGAA 720
DB 802 GAATACAGAGATGGCGGCCCAAAATTCACAGCCCTGTTATAGTGGAGAGTGCAGAA 861
QY 721 ATGATGTCAGAAATGGCTTCATGAGGATCAAAATGCTTCACTTAAGATGGGGCA 780
DB 862 ATGATGTCAGAAATGGCTTCATGAGGATCAAAATGCTTCACTTAAGATGGGGCA 921
QY 781 AAGACTTACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAATAT 840
DB 922 AAGACTTACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAATAT 981
QY 841 CCAGAAAGGTTGATCTGTCAGTGAACATCTGACAGCTGGGATGTTGGGAGGGTCC 900
DB 982 CCAGAAAGGTTGATCTGTCAGTGAACATCTGACAGCTGGGATGTTGGGAGGGTCC 1041
QY 901 ATGATGAT 909
DB 1042 ATGATGAT 1050
RESULT 6
BX355940
LOCUS
DEFINITION
BX355940 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1006Y115 5-PRIME, mRNA sequence.
BX355940
VERSION
BX355940.1 GI:30384019
EST.
ACCESSION
BX355940.1 GI:30384019
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1038)
L1, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1006AB08QPI&cluster=4663.f. Contact :
Feng liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1006AB08QPI.
Location/Qualifiers
1. 1038
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1006Y115"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)"

primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 304 a 237 c 262 g 233 t 2 others

Query Match 59.6%; Score 843.8; DB 13; Length 1038;
Best Local Similarity 99.7%; Pred. No. 8.1e-236;
Matches 856; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 ATGAATTCCTTATCTTGGCACTTTTGGTGGTTCACCTTTATCCCTGCTCTGGG 60
180 ATGAATTCCTTATCTTGGCACTTTTGGTGGTTCACCTTTATCCCTGCTCTGGG 239
61 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 120
240 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 299
121 GCCAGCTGTGAGATGTTGCTTAAGCAATCACTAGCTGTTTATGTAAGCCGAG 180
300 GCCAGCTGTGAGATGTTGCTTAAGCAATCACTAGCTGTTTATGTAAGCCGAG 359
181 AACAGATCTTATGAGCGATGGCACTTCTGTTGATCTGTTGAACCCAGATGAGTGC 240
360 AACAGATCTTATGAGCGATGGCACTTCTGTTGATCTGTTGAACCCAGATGAGTGC 419
241 TCCAGAACTTATGAGCGATGGCACTTCTGTTGATCTGTTGAACCCAGATGAGTGC 300
420 TCCAGAACTTATGAGCGATGGCACTTCTGTTGATCTGTTGAACCCAGATGAGTGC 479
301 GAGAAAGTTTCACTGAGAGCCAGTGAATATACCCCACTGGGAGAGGGAGAAATCAGCT 360
480 GAGAAAGTTTCACTGAGAGCCAGTGAATATACCCCACTGGGAGAGGGAGAAATCAGCT 539
361 GTGATCTGAGAGCCAGTGAATATACCCCACTGGGAGAGGGAGAAATCAGCT 420
540 GTGATCTGAGAGCCAGTGAATATACCCCACTGGGAGAGGGAGAAATCAGCT 599
421 ACTCTCCAGAAAGGATTAAGCAAGATTTCTGTTGATCTGTTGAACCCAGATGAGTGC 480
600 ACTCTCCAGAAAGGATTAAGCAAGATTTCTGTTGATCTGTTGAACCCAGATGAGTGC 659
481 AGAAGGGGCTCAGAGAGCAAGAGGAGATTTGTTTAAACCACTTATCACTCACTAC 540
660 AGAAGGGGCTCAGAGAGCAAGAGGAGATTTGTTTAAACCACTTATCACTCACTAC 719
541 TCAAGACGGTGCATTAACCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
720 TCAAGACGGTGCATTAACCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 779
601 GCATCTCTATTCGATCCGTTGGGCTCTTCTTCATCTTACAGTCTTCAACAGGATTTAG 660
780 GCATCTCTATTCGATCCGTTGGGCTCTTCTTCATCTTACAGTCTTCAACAGGATTTAG 839
661 GAATTCAGAGATGGGCTGCCCAAAATTCGAACAGCTGATATTCAGTGGAGATCAGAA 720
840 GAATTCAGAGATGGGCTGCCCAAAATTCGAACAGCTGATATTCAGTGGAGATCAGAA 899
721 ATGATGTCAAGAAATGGCTTCTGATGGATCAAAATTTGATTCAGCTTAAAGAT-GGGGGC 779
900 ATGATGTCAAGAAATGGCTTCTGATGGATCAAAATTTGATTCAGCTTAAAGATGGGGGC 959
780 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 839
960 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 1019
840 TCCAGAACTTATGAGCGATGGCACTTCTGTTGATCTGTTGAACCCAGATGAGTGC 240
1020 TCCAGAACTTATGAGCGATGGCACTTCTGTTGATCTGTTGAACCCAGATGAGTGC 300

RESULT 7
B0878966

LOCUS B0878966 909 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT 8183107 Lupski dorsal root ganglion Homo sapiens cDNA
clone IMAGE:6184129 5', mRNA sequence.
ACCESSION B0878966
VERSION B0878966.1 GI:22270974
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strauberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM13572 row: n column: 02
High quality sequence stop: 696.
Location/Qualifiers
1..909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184129"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_id="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCAGCCAGGCGTCCG-3' and 5'-GACTAGTTCATGATGCGAGGCGGCCCTT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 273 a 202 c 226 g 207 t 1 others

Query Match 56.3%; Score 796.8; DB 13; Length 909;
Best Local Similarity 98.8%; Pred. No. 4.8e-222;
Matches 823; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

1 ATGAATTCCTTATCTTGGCACTTTTGGTGGTTCACCTTTATCCCTGCTCTGGG 60
64 ATGAATTCCTTATCTTGGCACTTTTGGTGGTTCACCTTTATCCCTGCTCTGGG 123
61 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 120
124 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 183
121 GCCAGCTGTGAGATGTTGCTTAAGCAATCACTAGCTGTTTATGTAAGCCGAG 180
184 GCCAGCTGTGAGATGTTGCTTAAGCAATCACTAGCTGTTTATGTAAGCCGAG 243
244 AACAGATCTTATGAGCGATGGCACTTCTGTTGATCTGTTGAACCCAGATGAGTGC 303
241 TCCAGAACTTATGAGCGATGGCACTTCTGTTGATCTGTTGAACCCAGATGAGTGC 300
304 TCCAGAACTTATGAGCGATGGCACTTCTGTTGATCTGTTGAACCCAGATGAGTGC 363

```

QY 301 GAGAAATTCACTGAGCCAGTGAATATCCCACTGGAGAGAGGAGAAATCAAGCT 360
Db 364 GAGAAATTCACTGAGCCAGTGAATATCCCACTGGAGAGAGGAGAAATCAAGCT 423
QY 361 GTGATGTGAGAGCCAAATTCATTAAGTGGCCATCTGGTGTCTTGGACAGATTTGG 420
Db 424 GTGATGTGAGAGCCAAATTCATTAAGTGGCCATCTGGTGTCTTGGACAGATTTGG 483
QY 421 ACTCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGTGACCTTTTCATGAACTGAG 480
Db 484 ACTCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGTGACCTTTTCATGAACTGAG 543
QY 481 AGAAGGCTCTCAGAAAGCAGAGGAGAGATTTGTTTAAACCAACTTACATCACTAC 540
Db 544 AGAAGGCTCTCAGAAAGCAGAGGAGAGATTTGTTTAAACCAACTTACATCACTAC 603
QY 541 TCAAGAGCGGTGCAATACCAAGCCAGAGGGGGGGTGGAGCTGCCAAGGTGGGGCTTTG 600
Db 604 TCAAGAGCGGTGCAATACCAAGCCAGAGGGGGGGTGGAGCTGCCAAGGTGGGGCTTTG 663
QY 601 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCTCAGACAGGATTCAG 660
Db 664 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCTCAGACAGGATTCAG 723
QY 661 GAATACAGAGATGGCGTCCCAAAATTCACACAGCTGTATTAAGTGAAGATGCAGAA 720
Db 724 GAATACAGAGATGGCGTCCCAAAATTCACACAGCTGTATTAAGTGAAGATGCAGAA 783
QY 721 ATGATGTCAAGATGGCTTTCTCATGGATCAAAATTTGTCATTGACTTAAAGAT- 779
Db 784 ATGATGTCAAGATGGCTTTCTCATGGATCAAAATTTGTCATTGACTTAAAGATGGGGG 843
QY 780 AAAGACTTACCCAGATCTAGTCTTCAACACT-GTAGCAGAGATCACTGGG 831
Db 844 AAAGACTTACCCAGATCTAGTCTTCAACACTGTAGCAGAGATCACTGGG 896

```

```

RESULT 8
BX360507 1126 bp mRNA linear EST 05-MAY-2003
LOCUS BX360507 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1071A19 5-PRIME, mRNA sequence.
ACCESSION BX360507
VERSION BX360507.1 GI:30374439
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1126)
LJ.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library: seqref@genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1071A1900p1&cluster=4663.f. Contact:
Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0D1071A1900p1.
location/Qualifiers
1. 1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1071A19"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"

```

FEATURES

source

```

1. 1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1071A19"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"

```

```

/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 311 a 249 c 298 g 246 t 22 others
ORIGIN

```

```

Query Match 54.7%; Score 774.4; DB 13; Length 1126;
Best Local Similarity 96.6%; Pred. No. 2e-215;
Matches 833; Conservative 9; Mismatches 15; Indels 5; Gaps 5;

```

```

QY 1 ATGAATTCCTTATCTTGGCAATTTTGGTGTGTTCACCTTTATCCCTGTGCTGGG 60
Db 210 ATGAATTCCTTATCTTGGCAATTTTGGTGTGTTCACCTTTATCCCTGTGCTGGG 269
QY 61 AAGCTATATGCAAGATGGCATCTCTAAGAGACCTTTGAAGAAATTAAGAAATA 120
Db 270 AAGCTATATGCAAGATGGCATCTCTAAGAGACCTTTGAAGAAATTAAGAAATA 329
QY 121 GCCAGCTGTGAGATGTGTGTAAGCAATCATCACTAGCTGTTATGTTAAGCCAG 180
Db 330 GCCAGCTGTGAGATGTGTGTAAGCAATCATCACTAGCTGTTATGTTAAGCCAG 389
QY 181 AACAGATCTATAGGCGATTGGCACTTGTGTGATTAAGTGTGACCCAGACTGAGTGGC 240
Db 390 AACAGATCTATAGGCGATTGGCACTTGTGTGATTAAGTGTGACCCAGACTGAGTGGC 449
QY 241 TCCAGAACCTTAAGAAAGGATCCCAATTTATGACCAAACTGACGACAGATGGCTG 300
Db 450 TCCAGAACCTTAAGAAAGGATCCCAATTTATGACCAAACTGACGACAGATGGCTG 509
QY 301 GAGAAAGTTCACTGGAGCGAGTGAATATCCCACTGGAGAGGGAGAAATCAAGCT 360
Db 510 GAGAAAGTTCACTGGAGCGAGTGAATATCCCACTGGAGAGGGAGAAATCAAGCT 569
QY 361 GTGATGTGAGAGCAAGATTCATTAAGATAGCCATCTGGGCTTGGACAGACATTTGG 420
Db 570 GTGATGTGAGAGCAAGATTCATTAAGATAGCCATCTGGGCTTGGACAGACATTTGG 629
QY 421 ACTCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGTGACCTTTTCATGAACTGAG 480
Db 630 ACTCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGTGACCTTTTCATGAACTGAG 689
QY 481 AGAAGGCTCTCAGAAAGCAGAGGAGATTTGTTTAAACCAACTTACATCACTAC 540
Db 690 AGAAGGCTCTCAGAAAGCAGAGGAGATTTGTTTAAACCAACTTACATCACTAC 749
QY 541 TCAAGAGCGGTGCAATACCAAGCCAGAGGGGGGGTGGAGCTGCCAAGGTGGGGCTTTG 600
Db 750 TCAAGAGCGGTGCAATACCAAGCCAGAGGGGGGGTGGAGCTGCCAAGGTGGGGCTTTG 809
QY 601 GCATCTCTCATTTGATCCGTGGCTCTCTTCATCTACAGTCTCTCAGACAGGATTTGAG 660
Db 810 GCATCTCTCATTTGATCCGTGGCTCTCTTCATCTACAGTCTCTCAGACAGGATTTGAG 869
QY 661 GAATACAGAGATGGCGTCCCAAAATTCACAGCCGTGTTTACCGT-GGAAGATGAGA 719
Db 870 GAATACAGAGATGGCGTCCCAAAATTCACAGCCGTGTTTACCGT-GGAAGATGAGA 929
QY 720 AATGATGTCAAGATGGCTTTCTCA-TGGATCAAAATTTGTCATTGAGCTTAAAGTGGGG 778
Db 930 AATGATGTCAAGATGGCTTTCTCA-TGGATCAAAATTTGTCATTGAGCTTAAAGTGGGG 989
QY 779 CAAGAAGCTTACCAAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAT 838
Db 990 S-AARAMCTACCAAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAMA 1046
QY 839 ATCCAGAACAGTTGATCTGCT 860
Db 1047 TCARMAAGGTATGATGATGAT 1068

```

RESULT 9

BX335995
 LOCUS BX335995 919 bp mRNA linear EST 02-MAY-2003
 DEFINITION BX335995 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1022YU16 5-PRIME, mRNA sequence.
 ACCESSION BX335995
 VERSION BX335995.1 GI:30339459
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 919)
 AUTHORS Li, W.B., Gruber, C., Jeesee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91060 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4663.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D1022DE08Qp1&cluster=4663.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1022DE08Qp1.
 Location/Qualifiers
 1. 919
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1022YU16"
 /rissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 273 a 206 c 226 g 208 t
 ORIGIN
 Query Match 54.1%; Score 766.2; DB 13; Length 919;
 Best Local Similarity 98.5%; Pred. No. 4.6e-213;
 Matches 779; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
 Oy 1 ATGAATTCCTTATCTTGCATTTTTCGCTGCTTCACTTTTATCCCTGCTCTGG 60
 Db 126 ATGAATTCCTTATCTTGCATTTTTCGCTGCTTCACTTTTATCCCTGCTCTGG 185
 Oy 61 AAAGTATATGCAAGAAATGGCATCTCTAAGAGAGACTTTTGAAGAAATGAAGAAATA 120
 Db 186 AAAGTATATGCAAGAAATGGCATCTCTAAGAGAGACTTTTGAAGAAATGAAGAAATA 245
 Oy 121 GCCAGCTGAGAGATGTTGTTAAACAATCACTAGCTGTTTATGTTAAAGCCAG 180
 Db 246 GCCAGCTGAGAGATGTTGTTAAACAATCACTAGCTGTTTATGTTAAAGCCAG 305
 Oy 181 AACAGATCTTATAGCGATTGGCACTTCTGTTGTAAGTCTTTGACCCAGACTGAGTGC 240
 Db 306 AACAGATCTTATAGCGATTGGCACTTCTGTTGTAAGTCTTTGACCCAGACTGAGTGC 365
 Oy 241 TCCAAAGACCTTAAGAAAGCAATCCCAATTAATGTCACAAAACCTGCACCAAGATGGGCTG 300
 Db 366 TCCAAAGACCTTAAGAAAGCAATCCCAATTAATGTCACAAAACCTGCACCAAGATGGGCTG 425
 Oy 301 GAGAAAGTTCACTGAGAGCCAGTGAATAATCCCACTGGAGAGAGGAGAAAGATCACT 360
 Db 426 GAGAAAGTTCACTGAGAGCCAGTGAATAATCCCACTGGAGAGAGGAGAAAGATCACT 485
 Oy 361 GTGATGCTGAGAGCAAGATTCATTAAGATAGCCATCTGGGTCTTTGGAGCAGCATTTGG 420
 Db 486 GTGATGCTGAGAGCAAGATTCATTAAGATAGCCATCTGGGTCTTTGGAGCAGCATTTGG 545

Oy 421 ACTCTCCAGAAAGCATTTACAGCAGAGATTCTGGTGTGACCTCTTTCAGTGAACCTCAG 480
 Db 546 ACTCTCCAGAAAGCATTTACAGCAGAGATTCTGGTGTGACCTCTTTCAGTGAACCTCAG 605
 Oy 481 AGAAGGCTTCAGAGAGCAAGAGAGATTGTTTATTAACAACCTTATCATCACTAC 540
 Db 606 AGAAGGCTTCAGAGAGCAAGAGAGATTGTTTATTAACAACCTTATCATCACTAC 665
 Oy 541 TCAAGAGCGTGCATTAATCCGACGAGGCGGTGGAACCTGCAAGGTGGGGCTTG 600
 Db 666 TCAAGAGCGTGCATTAATCCGACGAGGCGGTGGAACCTGCAAGGTGGGGCTTG 725
 Oy 601 GCATCTCATTTGCATCGTGGCTCTCTTCCATCTACAGTCCCTCACAGGATTTAC 660
 Db 726 GCATCTCATTTGCATCGTGGCTCTCTTCCATCTACAGTCCCTCACAGGATTTAC 785
 Oy 661 GAATACCAAGATGCGTGGCCCAAAATTCACAACGCTGTATTACGTTGAAATGCGAA 720
 Db 786 GAATACCAAGATGCGTGGCCCAAGATTCCAACGCTGTATTACGTTGAAATGCGAA 845
 Oy 721 ATGATGCAAGAAAGGCTTCTCAT -GGGATCAAAATGTGATTCAGTAAAGATGGGGGC 779
 Db 846 ATGATGCAAGAAAGGCTTCTCATGGGATCAAAATGTGATTCAGTAAAGATGGGGGC 905
 Oy 780 AAAGACCTACC 790
 Db 906 AAAGACCTACC 916
 RESULT 10
 BU146905 866 bp mRNA linear EST 03-SEP-2002
 LOCUS BU146905
 DEFINITION AGENCOURT 7978204 Lupski dorsal_root_ganglion Homo sapiens cDNA
 clone IMAGE:6185233 5', mRNA sequence.
 ACCESSION BU146905
 VERSION BU146905.1 GI:22660437
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 866)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LHAM13575 row: 1 column: 02
 High quality sequence stop: 731.
 Location/Qualifiers
 1. 866
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6185233"
 /sex="male"
 /rissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6 (Life Technologies). Site 1:
 /clone_lib="Lupski dorsal root ganglion"
 /note="2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGCGTCCG-3' and
 5'-GACTAGTTTAAATCGCAGGCGCCCTT (15)-3'. Size selected >

1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies.

BASE COUNT 257 a 182 c 233 g 191 t 3 others

Query Match 52.9%; Score 748.6; DB 13; Length 866;
Best Local Similarity 99.1%; Pred. No. 6.6e-208;
Matches 773; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

65 CTATATGCAAGATGCGATCTCTAAGAGACTTT-TGAGAATATMAAGAATAGCC 123
Db 1 CTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAATATMAAGAATAGCC 60
Qy 124 AGCTGTGAGATGTTGCTAANAGCATCATCACTAGCTGTTATGTAAAGCCAGAAC 183
Db 61 AGCTGTGAGATGTTGCTAANAGCATCATCACTAGCTGTTATGTAAAGCCAGAAC 120
Qy 184 AGATCTATGAGCATGTCGACCTCTGTGATATCTGTGAGACCCAGACTGAGTCC 243
Db 121 AGATCTATGAGCATGTCGACCTCTGTGATATCTGTGAGACCCAGACTGAGTCC 180
Qy 244 AAGAACCTAGAAAAAGCCATCCAAATTATGTACAAAACTGCAGCAAGATGGCTGGAG 303
Db 181 AAGAACCTAGAAAAAGCCATCCAAATTATGTACAAAACTGCAGCAAGATGGCTGGAG 240
Qy 304 AAGATTCACCTGAGAGCCAGTGAATATCCCACTGGAGAGAGGAGAAATCAAGCTGTG 363
Db 241 AAGATTCACCTGAGAGCCAGTGAATATCCCACTGGAGAGAGGAGAAATCAAGCTGTG 300
Qy 364 ATGTGTGAGAGCAAAATTCATATAGATGACCATCTGGGTCTGTGAGAGAGATGGAGCT 423
Db 301 ATGTGTGAGAGCAAAATTCATATAGATGACCATCTGGGTCTGTGAGAGAGATGGAGCT 360
Qy 424 CCTCTCAGAGGCAATTAACAGAAATGCTGTGAGTGAACCTCTTTCAGTGAATCTGAGAGA 483
Db 361 CCTCTCAGAGGCAATTAACAGAAATGCTGTGAGTGAACCTCTTTCAGTGAATCTGAGAGA 420
Qy 484 AGGGCTTCAGAGAGCAAGAGGAGATGTTGTTTAAACAACTTACATCAACTACTCA 543
Db 421 AGGGCTTCAGAGAGCAAGAGGAGATGTTGTTTAAACAACTTACATCAACTACTCA 480
Qy 544 AGAGAGGTGCAATACCAAGCAGAGGGGGCGGTGAGAGCTGCCAGTGGGGCTTTGGCA 603
Db 481 AGAGAGGTGCAATACCAAGCAGAGGGGGCGGTGAGAGCTGCCAGTGGGGCTTTGGCA 540
Qy 604 TCTCTCATTCGATCGGTGGCTCCTCTCCATCTACATGATCCACAGATTCAGGAA 663
Db 541 TCTCTCATTCGATCGGTGGCTCCTCTCCATCTACATGATCCACAGATTCAGGAA 600
Qy 664 TACAGAGATGCGTGGCCCAAAATTCACACAGCTGTATTAAGGTGAGATGACAGAAATG 723
Db 601 TACAGAGATGCGTGGCCCAAAATTCACACAGCTGTATTAAGGTGAGATGACAGAAATG 660
Qy 724 ATGTCAAGATGCGTCTTCATGAGGATCAAAATGTTCATTCAGCTTAAAGATGGGGCAAG 783
Db 661 ATGTCAAGATGCGTCTTCATGAGGATCAAAATGTTCATTCAGCTTAAAGATGGGGCAAG 720
Qy 784 ACTTACCAAGATGATGATCTCTTCAACAAT-CTACAGAGATCACTGGGAGCAAAATATCC 842
Db 721 ACTTACCAAGATGATGATCTCTTCAACAATCTGTGAGAGATCACTGGGAGCAAAATATTC 780

RESULT 11
BX416896 1012 bp mRNA linear EST 13-MAY-2003
LOCUS BX416896 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE001YF04
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX416896
VERSION BX416896.1 GI:30650313
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1012)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished

COMMENT

Contact: Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE001DC02QPIcluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DE001DC02QPI.

FEATURES
source

1..1012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE001YF04"
/issue_type="PLACENTA"
/clone_id="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-Oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 281 a 240 c 275 g 215 t 1 others

ORIGIN

Query Match 52.5%; Score 742.8; DB 13; Length 1012;
Best Local Similarity 99.6%; Pred. No. 3.5e-206;
Matches 755; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 ATGAAATTCCTTATCTTGGCAATTTTGGTGGTGTACCTTTATCTCTGTGCTGGG 60
Db 255 ATGAAATTCCTTATCTTGGCAATTTTGGTGGTGTACCTTTATCTCTGTGCTGGG 314
Qy 61 AAACCTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAAATMAAGAATA 120
Db 315 AAACCTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAAATMAAGAATA 374
Qy 121 GCCAGCTGTGAGATGTTGCTAAGCAATCATCAACCTAGCTGTTATGTAAAGCCGAG 180
Db 375 GCCAGCTGTGAGATGTTGCTAAGCAATCATCAACCTAGCTGTTATGTAAAGCCGAG 434
Qy 181 AACAGATCTATGAGCGATTTGGCACTTCTGGTGTATCTGTGAGACCCAGACTGAGTGC 240
Db 435 AACAGATCTATGAGCGATTTGGCACTTCTGGTGTATCTGTGAGACCCAGACTGAGTGC 494
Qy 241 TCACAAGAACTAGAAAAAGCCATCCAAATTAATGTACAAAACCTGCAGCAAGATGGGCTG 300
Db 495 TCACAAGAACTAGAAAAAGCCATCCAAATTAATGTACAAAACCTGCAGCAAGATGGGCTG 554
Qy 301 GAGAAAGTTTCACTGAGAGCCAGTGAATATCCCACTGGGAGAGAGGAGAAATCAAGCT 360
Db 555 GAGAAAGTTTCACTGAGAGCCAGTGAATATCCCACTGGGAGAGAGGAGAAATCAAGCT 614
Qy 361 GTGATGCTGAGCCAGAAATTCATATAGATGACATCTCTGGTCTTTGGCAGCAGATTTGGG 420
Db 615 GTGATGCTGAGCCAGAAATTCATATAGATGACATCTCTGGTCTTTGGCAGCAGATTTGGG 674
Qy 421 ACTCTCTCAGAGGCAATTAACAGCAAGATTTGGTGTGATCTTTTCAGTGAATCTGAG 480
Db 675 ACTCTCTCAGAGGCAATTAACAGCAAGATTTGGTGTGATCTTTTCAGTGAATCTGAG 734
Qy 481 AGAAGGGCTCAGAGAGCAAGAGGAGATGTTGTTTAAACAACTTACATCAACTAC 540

Db 735 AGAAGGCGCTCAGGAAGCAGAGGAAAGATTGTTTATACCAACTTACATCACTAC 794
 Oy 541 TCAAGACGCTGCAATACCAAGCAGAGGAGGAGGAGGAGCTGCAAGG-793GGGCTTT 599
 Db 795 TCAAGACGCTGCAATACCAAGCAGAGGAGGAGGAGGAGCTGCAAGG-793GGGCTTT 854
 Oy 600 GGCATCTCTCATTCGATCCGTGGCCCTCTTCTCATCTACAGTCTCTCAACAGTATTC 659
 Db 855 GGCATCTCTCATTCGATCCGTGGCCCTCTTCTCATCTACAGTCTCTCAACAGTATTC 914
 Oy 660 GGAATACCAAGATGCGCTGCCCAAAATTCACACGCTGTATTAAGGTGAAGATGCA 719
 Db 915 GGAATACCAAGATGCGCTGCCCAAAATTCACACGCTGTATTAAGGTGAAGATGCA 974
 Oy 720 AATGATGTCAAGATGCGCTTCTCATGGATCAAAATTTG 757
 Db 975 AATGATGTCAAGATGCGCTTCTCATGGATCAAAATTTG 1012

RESULT 12
 LOCUS BG751497 902 bp mRNA linear EST 15-MAY-2001
 DEFINITION 602730076P1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4873770 5', mRNA sequence.

ACCESSION BG751497 GI:14062150
 VERSION BG751497
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. (bases 1 to 902)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
 JOURNAL
 AUTHORS
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
 Plate: L10M1752 row: 9 column: 19
 High quality sequence stop: 817.
 Location/Qualifiers

FEATURES
 source 1..902

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_image="4873770"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 43"
 /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library." 1 others

BASE COUNT 246 a 200 c 251 g 204 t

Query Match 52.3%; Score 740; DB 10; Length 902;

Best Local Similarity 95.7%; Pred. No. 2,2e-205;

Matches 847; Conservative 0; Mismatches 26; Indels 12; Gaps 8;
 331 CCCCACTGGAGAGGAGGAAGATCACTGTGATGTGAGCCCAAGATTCAATAGATA 390
 |||||

Db 2 CCCCACTGGAGAGGAGGAAGATCACTGTGATGTGAGCCCAAGATTCAATAGATA 61
 Oy 391 GCCATCTGGGCTCTTGGCAGAGCATTTGGGCTCTTCAGAGGATTAACAGAGAGTT 450
 Db 62 GCCATCTGGGCTCTTGGCAGAGCATTTGGGCTCTTCAGAGGATTAACAGAGAGTT 121
 Oy 451 CTGTGTGTGACCTCTTTCATGAACTGCAAGAGAGGCTTCAGAGCAAGAGGAGATT 510
 Db 122 CTGTGTGTGACCTCTTTCATGAACTGCAAGAGAGGCTTCAGAGCAAGAGGAGATT 180
 Oy 511 GTTGTATTAACCACTTACATCACTCAAGAGAGGCTGCAATCCGACAGAGG 570
 Db 181 GTTGTATTAACCACTTACATCACTCAAGAGAGGCTGCAATCCGACAGAGG 240
 Oy 571 GCGGTGAAGCTGCAAGAGTGGGGCTTTGGGCTCTTCATTCGATCCGTGGCTCTTC 630
 Db 241 GCGGTGAAGCTGCAAGAGTGGGGCTTTGGGCTCTTCATTCGATCCGTGGCTCTTC 300
 Oy 631 TCCATCTACAGTCTCTCAACAGGATTCAGAGATACCAAGATGCGTGCCTCAAAATTC 690
 Db 301 TCCATCTACAGTCTCTCAACAGGATTCAGAGATACCAAGATGCGTGCCTCAAAATTC 360
 Oy 691 ACAGCTCTATTTACGCTGGAAGATGCAAGAAATGATGCAAGATGCTTCTCATGGATC 750
 Db 361 ACAGCTCTATTTACGCTGGAAGATGCAAGAAATGATGCAAGATGCTTCTCATGGATC 420
 Oy 751 AAATGTCTATTCAGTAAAGATGGGGGCAAGACCTACCAAGATCTGATTCCTTCAAC 810
 Db 421 AAATGTCTATTCAGTAAAGATGGGGGCAAGACCTACCAAGATCTGATTCCTTCAAC 480
 Oy 811 ACTGTACAGAGATCACTGGGAGCAATATTCAGAACAGGTTGATCTGATGAGCAT 870
 Db 481 ACTGTACAGAGATCACTGGGAGCAATATTCAGAACAGGTTGATCTGATGAGCAT 540
 Oy 871 CTGACACGCTGGAGATGTTGGGCAAGGCTGCATGG-ATGATGGGCTGTGAGCTTTATATC 929
 Db 541 CTGACACGCTGGAGATGTTGGGCAAGGCTGCATGG-ATGATGGGCTGTGAGCTTTATATC 600
 Oy 930 ATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCAAGAGGACTGCGGCT 989
 Db 601 ATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCAAGAGGACTGCGGCT 660
 Oy 990 GGTGCTCTGACT-GCAGAAAGCAAGAGTGTGCTTCCAGATATTCAGTTAC 1048
 Db 661 GGTGCTCTGACTGCGAGAAAGCAAGAGTGTGCTTCCAGATATTCAGTTAC 720
 Oy 1049 ACAAGTAAATATTTCCACTA--CACTGTGTGATGAGTCTGACGCAAGAACTTCT 1105
 Db 721 ACAAGTAAATATTTCCACTAAGTCTGTGTGATGAGTCTTGAACGCAAGAACTTCT 780
 Oy 1106 T---ACCCACGTGGGCTGCAATTCAC--TGGCAGTGAAGAAAGGCGCATCATGAGAGAG 1161
 Db 781 TTAACCACTTGGGCTGCAATTCACCTTGGCAGTGAAGAAAGGCGCATCATGAGAGAG 840
 Oy 1162 GTTATGAGCTGTGACAGCCCTCAATATCACTCAGGCTCTGAGC 1206
 Db 841 GTTATGAGCTGTGAGAGCC--TCATATCACTCAGGCTCGAACC 883

RESULT 13
 LOCUS BX460463 998 bp mRNA linear EST 22-MAY-2003
 DEFINITION BX460463 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone CS0DP014YN05 5-PRIME, mRNA sequence.

ACCESSION BX460463 GI:31029218
 VERSION BX460463
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. (bases 1 to 998)
 L1, W. B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Genoscope -- Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODF014CG03QPLcluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSODF014CG03QPL.

FEATURES
SOURCE
1. 998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODF014YN05"
/issue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoR sites of the pCMVSPORT 6
vector. Library was not normalized."

BASE COUNT 280 a 232 c 233 g 223 t 30 others

ORIGIN

Query Match 52.0%; Score 736.6; DB 13; Length 998;
Best Local Similarity 97.4%; Pred. No. 2,3e-204; Indels 0; Gaps 0;
Matches 736; Conservative 9; Mismatches 11;

QY 1 ATGAATTCCTTATCTTGCATTTTTCGAGTGTTCACCTTTATCCCTGCTGGG 60
DB 239 ATGAATTCCTTATCTTGCATTTTTCGAGTGTTCACCTTTATCCCTGCTGGG 298

QY 61 AAAGCTATTCAGAGATGCGATCTTAAAGGACCTTTGAAGAAATTAAGAAATA 120
DB 299 AAAGCTATTCAGAGATGCGATCTTAAAGGACCTTTGAAGAAATTAAGAAATA 358

QY 121 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTTTATGTAAAGCCAG 180
DB 359 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTTTATGTAAAGCCAG 418

QY 181 AACAGATCTATGAGCGATTGGCACTTCTGTTGATCTTTGAGACCCAGCTGAGGC 240
DB 419 AACAGATCTATGAGCGATTGGCACTTCTGTTGATCTTTGAGACCCAGCTGAGGC 478

QY 241 TCCAGAGACCTTAAGAAAGCATTCATTAATTAATTAATTAATTAATTAATTA 300
DB 479 TCCAGAGACCTTAAGAAAGCATTCATTAATTAATTAATTAATTAATTAATTA 538

QY 301 GAGAAAGTTCACCTGAGCGAGTGAATACCCAGTGGGAGAGGAGAGAAATCAGCT 360
DB 539 GAGAAAGTTCACCTGAGCGAGTGAATACCCAGTGGGAGAGGAGAGAAATCAGCT 598

QY 361 GTGATGCTGAGCGAGAAATCATTAAGATAGCAATCTGCTGTTGAGCAGAGATTGGG 420
DB 599 GTGATGCTGAGCGAGAAATCATTAAGATAGCAATCTGCTGTTGAGCAGAGATTGGG 658

QY 421 ACTGCTCCAGAGGATTAACGACGAGAGTGTGCTGAGCTCTTTGATGAATCTGAG 480
DB 659 ACTGCTCCAGAGGATTAACGACGAGAGTGTGCTGAGCTCTTTGATGAATCTGAG 718

QY 481 AGAAGGCTCTAGAGAGAGAGAGAGTGTGTTTATTAACCACTTATTAACATCAG 540
DB 719 AGAAGGCTCTAGAGAGAGAGAGAGTGTGTTTATTAACCACTTATTAACATCAG 778

QY 541 TCAAGACGCTGCAATACCAAGCGAGGGGCGGTGAAGCTGCGCAAGGTGGGGCTTTG 600

DB 779 TCAAGACGCTGCAATACCAAGCGAGGGGCGGTGAAGCTGCGCAAGGTGGGGCTTTG 838
QY 601 GCATCTCTATTCGATCCGTCGCTCTCTTCATCTACAGTCTCCACAGGATTCAG 660
DB 839 GCATCTCTATTCGATCCGTCGCTCTCTTCATCTACAGTCTCCACAGGATTCAG 898

QY 661 GAATACCGAGATGCGCGCCCAAAATTCACAGCGCTGTATTAACGTGAAGATGAGAA 720
DB 899 GAATACCGAGATGCGCGCCCAAAATTCACAGCGCTGTATTAACGTGAAGATGAGAA 958

QY 721 ATGATGTCAAGATGCTTCTCATGGGATCAAAAT 756
DB 959 ATGATGTCAAGATGCTTCTCATGGGATCAAAAT 994

RESULT 14
BX355939/c 996 bp mRNA linear EST 05-MAY-2003
LOCUS
DEFINITION BX355939 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSODI006Y115 3-PRIME, mRNA sequence.
ACCESSION BX355939
VERSION BX355939.1 GI:30382027
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Genoscope -- Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODI006AE08NP1cluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSODI006AE08NP1.

FEATURES
SOURCE
1. 996
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI006Y115"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 270 a 231 c 223 g 269 t 3 others

ORIGIN

Query Match 51.3%; Score 726.6; DB 13; Length 996;
Best Local Similarity 99.9%; Pred. No. 2e-201; Indels 0; Gaps 0;
Matches 726; Conservative 1; Mismatches 0;

QY 690 AACAGCTGTATTAACGCTGAGAGATGCAAGATGATCAAGATGCTTCTCATGGGAT 749
DB 996 AACAGCTGTATTAACGCTGAGAGATGCAAGATGATCAAGATGCTTCTCATGGGAT 937

QY 750 CAATATGTCTATTCAGCTTAAGATGGGGCAAGAGCTTACCCAGATCTGATTCCTTCAA 809
DB 936 CAATATGTCTATTCAGCTTAAGATGGGGCAAGAGCTTACCCAGATCTGATTCCTTCAA 877

QY 810 CACTGTAGCAGAGATCACTGGAGCAATATTCAGAAACAGTTGTACTGTCTAGTGACA 869
DB 876 CACTGTAGCAGAGATCACTGGAGCAATATTCAGAAACAGTTGTACTGTCTAGTGACA 817


```

Oy 870 TCTGACAGCTGGAGTGTGGCAGGAGTCCATGATGATGGGAGAGCTTTATATC 929
Db 816 TCTGACAGCTGGAGTGTGGCAGGAGTCCATGATGATGGGAGAGCTTTATATC 757
Oy 930 ATGGAGACACTCTCACTTATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTTGCGCT 989
Db 756 ATGGAGACACTCTCACTTATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTTGCGCT 697
Oy 990 GGTGCTCTGAGCTGAGAAAGAAAGAGTGGATGGTGGCTTCCAGTATTAATCACTTACA 1049
Db 636 GGTGCTCTGAGCTGAGAAAGAAAGAGTGGATGGTGGCTTCCAGTATTAATCACTTACA 637
Oy 1050 CAAGGTAATATATTTCCAACTACAGTCTGGTATGAGTCTGACGAGAGACCTTCTTACC 1109
Db 636 CAAGGTAATATATTTCCAACTACAGTCTGGTATGAGTCTGACGAGAGACCTTCTTACC 577
Oy 1110 CACTGGGCTGCATTCATCTGAGAGTGAAGAAAGGCGCAGGCGCATCATGAGAGATTATGAG 1169
Db 576 CACTGGGCTGCATTCATCTGAGAGTGAAGAAAGGCGCAGGCGCATCATGAGAGATTATGAG 517
Oy 1170 CTTGCTGACAGCCCTTCATATATCACTGAGTCTTGAGCCATGAGAAAGGAGACATCA 1229
Db 516 CTTGCTGACAGCCCTTCATATATCACTGAGTCTTGAGCCATGAGAAAGGAGACATCA 457
Oy 1230 CTTTGGATCCAGGCTGAGGCTGAGGCGCATCTGATGATGATCACTTATCAAGTATTT 1289
Db 456 CTTTGGATCCAGGCTGAGGCTGAGGCGCATCTGATGATGATCACTTATCAAGTATTT 397
Oy 1290 CTTTTCATCACTCCACAGGAGACCACTGATCTGATGATCCAAAGAGATGATGT 1349
Db 396 CTTTTCATCACTCCACAGGAGACCACTGATCTGATGATCCAAAGAGATGATGT 337
Oy 1350 TGTCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTGACACATGAGAAAGATCTGCC 1409
Db 336 TGTCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTGACACATGAGAAAGATCTGCC 277
Oy 1410 TAGGTCC 1416
Db 276 TAGGTCC 270

```

```

RESULT 15
CB958693 739 bp mRNA linear EST 29-APR-2003
LOCUS AGENCOURT.13778385 NIH_MGC.184 Homo sapiens cDNA clone
DEFINITION IMAGE:30351744 5', mRNA sequence.
ACCESSION CB958693
VERSION CB958693.1 GI:30214809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 739)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palokovics
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM149 row: m column: 01
High quality sequence shop: 572.
Location/Qualifiers
1..739
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db xref="taxon:9606"
/clone="IMAGE:30351744"
/lab host="DH10B (TI, phage-resistant)"
/clone_lib="NIH_MGC_184"
/note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site_1:
Site1 (ggccatctatggcc); Site_2: Site1 (ggccctccggcc);
Library is oligo-dT primed and directionally cloned. cDNA
was prepared from a glandular pool of tissues from thyroid,
parathyroid, adrenal, cortex and pineal gland. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTTATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCGCGCATG-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.38
kb (range 0.60-3.5 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."
BASE COUNT 192 a 171 c 195 g 181 t
ORIGIN
Query Match 50.8%; Score 719.2; DB 14; Length 739;
Best local similarity 99.6%; Pred. No. 2.6e-199;
Matches 721; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 588 GGTGGGGGCTTTGGCATCTCTCATTCGATCGTGCGCTCTTCATCTACAGTCTCA 647
Db 3 GGTGGGGGCTTTGGCATCTCTCATTCGATCGTGCGCTCTTCATCTACAGTCTCA 62
Oy 648 CACAGTATTCAGGAATACAGAGATGCGTCCCAAAATTCAGAGCTGTATACGCT 707
Db 63 CACAGTATTCAGGAATACAGAGATGCGTCCCAAAATTCAGAGCTGTATACGCT 122
Oy 708 GGAAGATGCGAAGATGATGTCAGAAATGGCTTCATGGGATCAAAATGTCAATGACT 767
Db 123 GGAAGATGCGAAGATGATGTCAGAAATGGCTTCATGGGATCAAAATGTCAATGACT 182
Oy 768 AAGATGGGGGCAAGAGACTACCCAGATGATGATTCCTTCAACACTGAGAGATCAC 827
Db 183 AAGATGGGGGCAAGAGACTACCCAGATGATGATTCCTTCAACACTGAGAGATCAC 242
Oy 828 TGGAGCAAAATATCAAGAACAGTGTGATCTGTGACAGTGAACATTTGAGACGCTGGATGT 887
Db 243 TGGAGCAAAATATCAAGAACAGTGTGATCTGTGACAGTGAACATTTGAGACGCTGGATGT 302
Oy 888 TGGGCAAGGTCGCAATGATGATGCGGCTGGAGAGCTTTATATCAATGAGAGACTCTCACT 947
Db 303 TGGGCAAGGTCGCAATGATGATGCGGCTGGAGAGCTTTATATCAATGAGAGACTCTCACT 362
Oy 948 TATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTGCGCTGTGCTCTGAGCTGAGAGA 1007
Db 363 TATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTGCGCTGTGCTCTGAGCTGAGAGA 422
Oy 1008 AGAACAGTGGAGTGTGGCTTCCAGTATTAATCACTACACAGGTAATATTTCAA 1067
Db 423 AGAACAGTGGAGTGTGGCTTCCAGTATTAATCACTACACAGGTAATATTTCAA 482
Oy 1068 CTAAGAGCTGGTATGAGAGTCTGACGAGAACTTCTTCCCACTGGGCTGCAATTCAC 1127
Db 483 CTAAGAGCTGGTATGAGAGTCTGACGAGAACTTCTTCCCACTGGGCTGCAATTCAC 542
Oy 1128 TGGCAGTGAAGAGCCAGGCGCATCATGAGAGGTTATAGCTCTGAGAGCCCTCAA 1187
Db 543 TGGCAGTGAAGAGCCAGGCGCATCATGAGAGGTTATAGCTCTGAGAGCCCTCAA 602
Oy 1188 TATCACTCAGGCTCTGAGCCATGAGAGAGGACAGACATCACTTTGGATCCAGCTGG 1247
Db 603 TATCACTCAGGCTCTGAGCCATGAGAGAGGACAGACATCACTTTGGATCCAGCTGG 662
Oy 1248 AGTCCCTGAGAGCCAGTCTACTGATGATGATCAATATTTCTTCCATCACTCCA 1307
Db 663 AGTCCCTGAGAGCCAGTCTACTGATGATGATCAATATTTCTTCCATCACTCCA 722
Oy 1308 CGGA 1311

```


Db 723 CGGA 726

Search completed: December 22, 2003, 23:11:28
Job time : 2776.81 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:59:40 ; Search time 93.5249 Seconds
(without alignments)
6682.695 Million cell updates/sec

Title: US-09-745-763-35_COPY_99_1514
Perfect score: 1416
Sequence: 1 ATGAATTCCTATCTTCG.....AAGAATGCTGCTAGGTCC 1416

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1412.8	99.8	1863	US-09-482-273-28	Sequence 28, Appl
2	772.4	54.5	1134	US-09-482-273-95	Sequence 95, Appl
3	49.2	3.5	7218	US-08-232-463-14	Sequence 14, Appl
4	37	2.6	2049	US-08-268-797-1	Sequence 1, Appl
5	37	2.6	2049	PCT-US95-08414-1	Sequence 1, Appl
6	37	2.6	2691	US-07-878-960-1	Sequence 1, Appl
7	35.2	2.5	2373	US-08-980-080-1	Sequence 1, Appl
8	35.2	2.5	3003	US-08-434-730-15	Sequence 15, Appl
9	34.6	2.4	4086	US-09-220-132-56	Sequence 56, Appl
10	34.6	2.4	4334	US-09-620-312D-88	Sequence 88, Appl
11	34.4	2.4	738	US-08-738-462-1	Sequence 1, Appl
12	34.4	2.4	738	PCT-US94-07587-1	Sequence 1, Appl
13	34	2.4	2106	US-09-252-991A-7477	Sequence 7477, Ap
14	34	2.4	2112	US-09-252-991A-7334	Sequence 7334, Ap
15	34	2.4	2187	US-09-252-991A-7167	Sequence 7167, Ap
16	33.4	2.4	2201	US-09-330-970-2	Sequence 2, Appl
17	33.4	2.4	2201	US-09-330-970-4	Sequence 4, Appl
18	33.2	2.3	2443	US-08-745-934-2	Sequence 2, Appl
19	33.2	2.3	9763	US-08-973-273-1	Sequence 1, Appl
20	33	2.3	1165	US-09-023-942A-28	Sequence 28, Appl
21	33	2.3	1443	US-09-328-352-2934	Sequence 2934, Ap
22	33	2.3	4791	US-09-328-352-1416	Sequence 1416, Ap
23	32.4	2.3	3483	US-09-620-312D-658	Sequence 658, App
24	32.2	2.3	272	US-09-313-294A-3342	Sequence 3342, Ap
25	32.2	2.3	1650	US-09-220-132-103	Sequence 103, App
26	32.2	2.3	2220	US-08-997-251-1	Sequence 1, Appl
27	32	2.3	4631	US-09-620-312D-164	Sequence 164, App

28	31.8	2.2	615	US-09-107-532A-1080	Sequence 1080, Ap
29	31.6	2.2	949	US-09-221-017B-594	Sequence 594, App
c	30	31.6	28473	US-08-961-527-83	Sequence 83, Appl
31	31.4	2.2	2242	US-09-400-742-1	Sequence 1, Appl
32	31.4	2.2	2242	US-08-618-651A-1	Sequence 1, Appl
33	31.4	2.2	2242	US-09-215-252-1	Sequence 1, Appl
34	31.2	2.2	1712	US-08-809-326A-8	Sequence 8, Appl
35	31.2	2.2	1712	US-08-689-914A-8	Sequence 8, Appl
36	31.2	2.2	1712	US-09-689-913A-8	Sequence 8, Appl
37	31.2	2.2	1712	US-09-689-916A-8	Sequence 8, Appl
38	31.2	2.2	1927	US-09-689-914A-13	Sequence 13, Appl
39	31.2	2.2	1927	US-09-689-913A-13	Sequence 13, Appl
40	31.2	2.2	1927	US-09-689-916A-13	Sequence 13, Appl
41	31.2	2.2	1954	US-08-809-326A-13	Sequence 13, Appl
c	42	31	280	US-08-783-853A-34	Sequence 34, Appl
c	43	31	280	US-09-344-050-34	Sequence 34, Appl
c	44	31	321	US-08-783-853A-43	Sequence 43, Appl
c	45	31	321	US-09-344-050-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-09-482-273-28
Sequence 28, Application US/09482273.
Patent No. 6534631.
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
CURRENT FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 1863
TYPE: DNA
ORGANISM: Homo sapiens
US-09-482-273-28
Query Match 99.8%; Score 1412.8; DB 4; Length 1863;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
c
1 ATGAATTCCTATCTTCGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 60
|||
99 ATGAATTCCTATCTTCGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 158
c
61 AAAGCTATGCAAGATGCGATCTCTAGAGACCTTTGAAGAAATTAAGAATA 120
|||
159 AAAGCTATGCAAGATGCGATCTCTAGAGACCTTTGAAGAAATTAAGAATA 218
c
121 GCCAGCTGGAGATGTTGCTTAAGCAATCACTCAAGCTGTTTATGTAAGCCAG 180
|||
219 GCCAGCTGGAGATGTTGCTTAAGCAATCACTCAAGCTGTTTATGTAAGCCAG 278
c
181 AACAGATCTATGAGCGATGCGACTTCTGTTGATACCTGTTGAGACCGAGATGAGGC 240
|||
279 AACAGATCTATGAGCGATGCGACTTCTGTTGATACCTGTTGAGACCGAGATGAGGC 338
c
241 TCCAAGAACTGAAAGAAAGCCATCAATTAATGTAACAAACCTGCAAGAAATGAGGCTG 300
|||
339 TCCAAGAACTGAAAGAAAGCCATCAATTAATGTAACAAACCTGCAAGAAATGAGGCTG 398

301 GAGAAAGTTCACTGAGACAGTGAATACCCACTGAGAGAGGAGAAATCAAGCT 360
399 GAGAAAGTTCACTGAGACAGTGAATACCCACTGAGAGAGGAGAAATCAAGCT 458
361 GTGATGCTGAGACCAAGAAATTCATAGATACCATCTGGGTCTTGGCAGACATTTGGG 420
459 GTGATGCTGAGACCAAGAAATTCATAGATACCATCTGGGTCTTGGCAGACATTTGGG 518
421 ACTCTCCAGAAAGGATTCAGACAGAAATTCGTGGGTGAGACTCTTTGAGAACTGAG 480
519 ACTCTCCAGAAAGGATTCAGACAGAAATTCGTGGGTGAGACTCTTTGAGAACTGAG 578
481 AGAAGGCTCAGAAAGCAAGAGAGAGATGTTGTTTAAACCACTTACATCACTAC 540
579 AGAAGGCTCAGAAAGCAAGAGAGAGATGTTGTTTAAACCACTTACATCACTAC 638
541 TCAAGACGCTGCAATACCGAACGACGAGGCGGTGAGAGCTGCCAGGTGGGCTTTG 600
639 TCAAGACGCTGCAATACCGAACGACGAGGCGGTGAGAGCTGCCAGGTGGGCTTTG 698
601 GCATCTCTCATTCATCCGTGGCCCTCTCCATCTACAGTCCCTCAGACAGATTTGAG 660
699 GCATCTCTCATTCATCCGTGGCCCTCTCTCCATCTACAGTCCCTCAGACAGATTTGAG 758
661 GAATACCAAGATGGCGTCCCAAAATTCACACAGCTGTATTACGAGTGAAGATGAGAA 720
759 GAATACCAAGATGGCGTCCCAAAATTCACACAGCTGTATTACGAGTGAAGATGAGAA 818
721 ATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTATTCACTTAAAGATGGGCGCA 780
819 ATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTATTCACTTAAAGATGGGCGCA 878
781 AAGACCTACCAAGATGCTGATTCCTTCAACACTGAGAGAGATCACAGTGGAGCAAAATAT 840
875 AAGACCTACCAAGATGCTGATTCCTTCAACACTGAGAGAGATCACAGTGGAGCAAAATAT 938
841 CCAGAACAGATTTGATCTGATGATGACATCTGACACAGTGGAGATGGGCGGCTCC 900
935 CCAGAACAGATTTGATCTGATGATGACATCTGACACAGTGGAGATGGGCGGCTCC 998
901 ATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTATTCACTTAAAGATGGT 960
999 ATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTATTCACTTAAAGATGGT 1058
961 GGGGCTGCTCAAGAGAGACTCTGAGGCTGCTGAGACAGCAAGAAAGAGTGGGA 1020
1059 GGGGCTGCTCAAGAGAGACTCTGAGGCTGCTGAGACAGCAAGAAAGAGTGGGA 1118
1021 GTTGGTGGCTTCCAGATATTATCATGTTACACAGAGTAATTTCCACTACAGTCTGGTG 1080
1119 GTTGGTGGCTTCCAGATATTATCATGTTACACAGAGTAATTTCCACTACAGTCTGGTG 1178
1081 ATGAGATCTGACGACGAACTTCTTACCACTGGGCTGCAATTTCACTGGCAGTGAAG 1140
1179 ATGAGATCTGACGACGAACTTCTTACCACTGGGCTGCAATTTCACTGGCAGTGAAG 1238
1141 GCGCAGGCGCATCAGAGAGAGATTAAGCTGCTGACGCCCCCTCAATATCACTCAGGTC 1200
1239 GCGCAGGCGCATCAGAGAGAGATTAAGCTGCTGACGCCCCCTCAATATCACTCAGGTC 1298
1201 CTGAGCCATGAGAGAGAGAGAGATTAAGCTTGGATCCAAAGCTGAGAGCTCTGAGCC 1260
1299 CTGAGCCATGAGAGAGAGAGAGATTAAGCTTGGATCCAAAGCTGAGAGCTCTGAGCC 1358
1261 AGTCTACTGATGATCTTATACAGATTTCTTCTTCATCACTCCACGAGAGACATCAG 1320
1359 AGTCTACTGATGATCTTATACAGATTTCTTCTTCATCACTCCACGAGAGACATCAG 1418
1321 ACTGTCATGATCCAAAGAGAGATTAAGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1380
1419 ACTGTCATGATCCAAAGAGAGATTAAGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1478
1381 GTTGTTCAGACATGAGAGAAATGCTGCTAGTCC 1416

Db 1479 GTTGTTCAGACATGAGAGAAATGCTGCTAGTCC 1514
RESULT 2
US-09-482-273-95
; Sequence 95, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Roese et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; EARLIER FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-482-273-95
Query Match 54.5%; Score 772.4; DB 4; Length 1134;
Best Local Similarity 99.7%; Pred. No. 1e-233;
Matches 784; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 631 TCATCTACAGTCTCTCAACAGTATTCAGAAATCCAGATGCGTCCCAAAATTCGA 690
Db 1 TCATCTACAGTCTCTCAACAGTATTCAGAAATCCAGATGCGTCCCAAAATTCGA 60
QY 691 ACAGCCGTATTCAGGAGAGATTCAGAAATGATGCAAGATGCTTCATAGGATC 750
Db 61 ACAGCCGTATTCAGGAGAGATTCAGAAATGATGCAAGATGCTTCATAGGATC 120
QY 751 AAAATGTCATTCAGTAAAGATGGGGCAAGACCTACCACTGATTCCTTCAC 810
Db 121 AAAATGTCATTCAGTAAAGATGGGGCAAGACCTACCACTGATTCCTTCAC 180
QY 811 ACTGACAGAGATCACTGGAGCAATATTCAGAACAGTGTACTGTGACATGAT 870
Db 181 ACTGACAGAGATCACTGGAGCAATATTCAGAACAGTGTACTGTGACATGAT 240
QY 871 CTGAGACGCTGGAGATGTTGGGCAAGGCTGCATGATGATGATGAGCTTTATATCA 930
Db 241 CTGAGACGCTGGAGATGTTGGGCAAGGCTGCATGATGATGATGAGCTTTATATCA 300
QY 931 TGGGAGACCTCTCATCTTATTAAGATCTTGGCTGCTGCAAAAGAGACTTGGCGCTG 990
Db 301 TGGGAGACCTCTCATCTTATTAAGATCTTGGCTGCTGCAAAAGAGACTTGGCGCTG 360
QY 991 GTGCTGAGCTGAGAGAGAGAGAGAGTGTGCTTCCAGTATTCAGTTAC 1050
Db 361 GTGCTGAGCTGAGAGAGAGAGAGAGTGTGCTTCCAGTATTCAGTTAC 420
QY 1051 AAGGTAAATTTTCCACTCAGTCTGTGATGATGATGATGATGATGATGATGATGATGAT 1110
Db 421 AAGGTAAATTTTCCACTCAGTCTGTGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1111 ACTGGGCTGCAATTCATGAG 1170
Db 481 ACTGGGCTGCAATTCATGAG 539
QY 1171 CTGCTGAGCCCTCAATATCACTCAGGTCTGAGCCATGAGAGAGAGAGAGAGAGAGAGAG 1230
Db 540 CTGCTGAGCCCTCAATATCACTCAGGTCTGAGCCATGAGAGAGAGAGAGAGAGAGAGAG 599

Oy	1023	IGGAGCCCTTCAGATATATACAGTTACACAGATAATATTTCCACACCTACAGCTGTGAT	1082
Db	1252	AAGTCGCTTGAGTTCCTTTGTACCAACCAAGGACATCCGGCCCGCTCTTACATTCCTGT	1311
Oy	1083	GGAGTCTGACGAGGACCTTTCTTACCACTGGGCTGCATTCACCTGGCAGTGAAGAAAGG	1142
Db	1312	TTCTACTATTAATGATGTGTCGCACACCAAGCCCTGGAGATGCTGACAGGGAGAACAC	1371
Oy	1143	CAGGGCATCATGAGGAGG	1162
Db	1372	GTTTGTGTTAAGGTAGAAG	1391

RESULT 9

```

US-09-220-132-56/c
; Sequence 56, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 4086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-56

```

	Query Match	2.4%	Score 34.6	DB 4	Length 4086
	Best Local Similarity	56.6%	Pred. No. 2.7		
	Matches	64	Conservative	0	Gaps 0
Oy	1251	GCCTGAGCCACTGCTTGTGATGATTTATTTCTTCCATCACTCCACAG	1310		
Db	1678	GCATTACCCGCTGCTGCTGAACCTCAACTGCAGACATGTCATCTCCAGACTTCACCT	1619		
Oy	1311	AGACACCACTGTCATGATGATCAAGAAGATGATGTTGCTGCGCTTT	1363		
Db	1618	CGAGCTCTGTTCTTCCAGACCAACCTCTATGATATGATATTTTGTAT	1566		

RESULT 10

US-09-620-312D-88/C
Sequence 88, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aildong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunwei
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, No. 6569662el
TITLE OF INVENTION: Nucleic Acids and

```

? TITLE OF INVENTION: Polypeptides
? FILE REFERENCE: 784CIP2B
? CURRENT APPLICATION NUMBER: US/09/620,312D
? CURRENT FILING DATE: 2000-07-19
? PRIOR APPLICATION NUMBER: 09/552,317
? PRIOR FILING DATE: 2000-04-25
? PRIOR APPLICATION NUMBER: 09/468,725
? PRIOR FILING DATE: 2000-01-21
? NUMBER OF SEQ ID NOS: 1105
? SOFTWARE: Pf_Fl_genes Version 1.0
? SEQ ID NO 88
? LENGTH: 4334
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (64)..(3810)
US-09-620-312D-88

```

Query Match

	Query Match	Score	DB	Length
	Best Local Similarity	56.6%		Pred. No. 2.8;
Matches	64;	Conservative	0;	Mismatches 49; Indels 0; Gaps 0;
QY	1251	GCCTGGAGCCAGTCTACTGATGACTTATACAAAGATTTCTTCTTCATCTCACTCCACAG	1310	
DB	1915	GCATGTATCCCCGTCTGCTCTCACTCACTCACTGACACATCTATCTCTCCAGCACTTCCACCT	1855	
QY	1311	AGACACCATGACTGTGATGCATGCCAAGACAGATGAATGTTCTGCTGCTGTTT	1363	
DB	1855	CCAGGCTCTTGTTCCTTCAGACCAACATCTCATCATGATGATCTTATTTTGTAT	1803	

RESULT 11

US-08-738-462-1/c
Sequence 1, Application US/08738462
Patent No. 5965401
GENERAL INFORMATION:
APPLICANT: Chang, Chiwen
APPLICANT: Lanier, Lewis J.
APPLICANT: Phillips Jr., Joseph H.
TITLE OF INVENTION: Purified Mammalian NK Antigens and
TITLE OF INVENTION: Related Reagents
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: DMAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/738,462
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,435
FILING DATE: 16-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Chang, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0397
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-9196
TELEFAX: 415-496-1200
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 61..738
US-08-738-462-1

Query Match 2.4%; Score 34.4; DB 2; Length 738;
Best Local Similarity 48.4%; Pred. No. 1.1;
Matches 124; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 696 CTGTATTACGGTGAAGATGCAGAAATGATGCAGAAATGCTTCTCATGGGATCAAAAT 755
DB 261 CTGTATTAAAGATGTCTACTGAAACCTCAACCCAGTAAACAACAAGCAAGAGAAATAT 202
QY 756 TGTCACTTCACTTAAAGATGGGGCAAGACCTTACCAGAT-ACGTATCTCTCAACACTG 814
DB 201 CCCAGCACACTTAAGTTTCAGGGCAATTTGATGCCAAGGTGAAACCTGACAGACATCCG 142
QY 815 TAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTGATGAGATCTGG 874
DB 141 AGAAGAGATGAAGGTGAAGAACTTTCTGGGCTGAGTCTGTGGTTAAGTTTAACTAGC 82
QY 875 ACAGCTGGATGTTGGCAGGGTGCCATGATGATGCGCGTGGAGCCTTTATATCATGGG 934
DB 81 ATATATTGCTTGTGTCATGCGCAGACAGAGAGGTGGATTAACCTGTGTGAAGA 22
QY 935 AAGCACTCTCACTTAT 950
DB 21 ACAACTCTCAATTCT 6

RESULT 12

PCT-US94-07587-1/c
Sequence 1, Application PC/TUS9407587
GENERAL INFORMATION:
APPLICANT: Schering Corp.
TITLE OF INVENTION: PURIFIED MAMMALIAN NK ANTIGENS AND
TITLE OF INVENTION: RELATED REAGENTS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: One Giralda Farms
City: Madison
STATE: New Jersey
COUNTRY: USA
ZIP: 07940
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07587
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G.
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: DX0397K
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-822-7255
TELEFAX: 201-822-7039
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS

LOCATION: 61..738
PCT-US94-07587-1

Query Match 2.4%; Score 34.4; DB 5; Length 738;
Best Local Similarity 48.4%; Pred. No. 1.1;
Matches 124; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 696 CTGTATTACGGTGAAGATGCAGAAATGATGCAGAAATGCTTCTCATGGGATCAAAAT 755
DB 261 CTGTATTAAAGATGTCTACTGAAACCTCAACCCAGTAAACAACAAGCAAGAGAAATAT 202
QY 756 TGTCACTTCACTTAAAGATGGGGCAAGACCTTACCAGAT-ACGTATCTCTCAACACTG 814
DB 201 CCCAGCACACTTAAGTTTCAGGGCAATTTGATGCCAAGGTGAAACCTGACAGACATCCG 142
QY 815 TAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTGATGAGATCTGG 874
DB 141 AGAAGAGATGAAGGTGAAGAACTTTCTGGGCTGAGTCTGTGGTTAAGTTTAACTAGC 82
QY 875 ACAGCTGGATGTTGGCAGGGTGCCATGATGATGCGCGTGGAGCCTTTATATCATGGG 934
DB 81 ATATATTGCTTGTGTCATGCGCAGACAGAGAGGTGGATTAACCTGTGTGAAGA 22
QY 935 AAGCACTCTCACTTAT 950
DB 21 ACAACTCTCAATTCT 6

RESULT 13

US-09-252-991A-7477
Sequence 7477, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7477
LENGTH: 2106
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7477

Query Match 2.4%; Score 34; DB 4; Length 2106;
Best Local Similarity 54.9%; Pred. No. 2.8;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1042 CAGTTACACAAGGTAATATTTCCAACTACAGTCTGGTGAAGTGTGACGAGAAAC 1101
DB 1181 CAGGAAGTGAAGACCTGATGCTCTACTTCGCGCTGTGTGAACCCGACGACGACAC 1240
QY 1102 TTCTTACCACTGGGCTGCAATTCATCTGCGAGTGAAGGACGAGGCGCATCATGAGAG 1161
DB 1241 GCCTTCCTCAGGGTGAATCAACGTCGCGCGGGAATCGTTTCGGAGACCTGGAAG 1300
QY 1162 GT 1163
DB 1301 CT 1302

RESULT 14

US-09-252-991A-7334
Sequence 7334, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

Job time : 95.5249 secs

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7334
LENGTH: 2112
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7334

Query Match 2.4%; Score 34; DB 4; Length 2112;
Best Local Similarity 54.9%; Pred. No. 2.8;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1042 CAGTTACACAGGTAAATATTTCCAACTACAGTCTGTGATGAGTCTGACGAGAAC 1101
DB 1243 CAGGAAGTGAAGACCTGATGCTTCTACTTCGCGCTGTGTGAACCCGACGACGAC 1302
QY 1102 TTCTTACCCACTGGGCTGCAATTCACTGCGAGTGAAGGCCAGGCCCATCATGAGGAG 1161
DB 1303 GCCTTCCTCAGGGTATACAGCTGCGCGCGGGAATCGGTTCCGCGACCTTGAGAG 1362
QY 1162 GT 1163
DB 1363 CT 1364

RESULT 15
US-09-252-991A-7167/c
Sequence 7167, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7167
LENGTH: 2187
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7167

Query Match 2.4%; Score 34; DB 4; Length 2187;
Best Local Similarity 54.9%; Pred. No. 2.9;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1042 CAGTTACACAGGTAAATATTTCCAACTACAGTCTGTGATGAGTCTGACGAGAAC 1101
DB 975 CAGGAAGTGAAGACCTGATGCTTCTACTTCGCGCTGTGTGAACCCGACGACGAC 916
QY 1102 TTCTTACCCACTGGGCTGCAATTCACTGCGAGTGAAGGCCAGGCCCATCATGAGGAG 1161
DB 915 GCCTTCCTCAGGGTATACAGCTGCGCGCGGGAATCGGTTCCGCGACCTTGAGAG 856
QY 1162 GT 1163
DB 855 CT 854

Search completed: December 22, 2003, 23:17:19

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 14:49:05 ; Search time 443.013 Seconds
(without alignments)
10650.710 Million cell updates/sec

Title: US-09-745-763-35_COPY_99_1514
Perfect score: 1416
Sequence: 1 ATGAAATTCCTATCTTCG.....AGAAATGCTGCTAGTCC 1416

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2211978 seqs, 1666101734 residues

Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1416	100.0	1851	9	US-09-745-763-35
2	1412.8	99.8	1863	11	US-09-984-271-28
3	1107.2	78.2	1778	10	US-09-917-800A-505
4	772.4	54.5	1134	11	US-09-984-271-95
5	261.8	18.5	357	10	US-09-833-381-1929
6	246.2	17.4	427	10	US-09-833-381-1930
7	91	6.4	134	10	US-09-783-590-12141
8	60.2	4.3	65	13	US-09-908-975-25931
9	40.2	4.2	60	13	US-09-908-975-8960
10	42	3.0	2200	15	US-10-059-585-5
11	41	2.9	65	13	US-09-908-975-1202
12	39.4	2.8	484	11	US-09-918-995-30889
13	39.4	2.8	1363	15	US-10-232-484-5
14	39.4	2.8	2319	14	US-10-044-090-579
15	39.4	2.8	3322	13	US-09-814-353-20277

C 16	38.6	2.7	3997	15	US-10-198-846-12639	Sequence 12639, A
17	38	2.7	457	13	US-10-027-633-310889	Sequence 310889,
18	38	2.7	457	13	US-10-027-633-310889	Sequence 310889,
19	37.8	2.7	999	15	US-10-184-644-434	Sequence 434, App
20	37.8	2.7	999	15	US-10-184-644-434	Sequence 434, App
21	37.8	2.6	502	15	US-10-066-543-2028	Sequence 2028, Ap
22	37	2.6	548	15	US-10-066-543-2222	Sequence 2222, Ap
23	37	2.6	568	10	US-09-998-595-2545	Sequence 2545, Ap
24	37	2.6	2691	9	US-09-922-217-121	Sequence 121, App
25	37	2.6	2691	10	US-09-833-263-111	Sequence 121, App
26	37	2.6	2691	10	US-09-880-107-2399	Sequence 2399, Ap
27	37	2.6	2691	13	US-10-301-822-200	Sequence 200, App
28	37	2.6	2691	13	US-09-873-367C-1009	Sequence 1009, Ap
29	37	2.6	2691	14	US-10-025-380-121	Sequence 121, App
30	37	2.6	2691	15	US-10-102-524-1708	Sequence 1708, Ap
31	37	2.6	2782	14	US-10-044-090-647	Sequence 647, App
32	37	2.6	2782	15	US-10-208-408-16	Sequence 16, App1
33	37	2.6	2782	15	US-10-084-817-28	Sequence 28, App1
34	37	2.6	3313	13	US-09-814-353-19947	Sequence 19947, A
35	37	2.6	3320	15	US-10-198-846-11549	Sequence 11549, A
36	36.4	2.6	5085	15	US-10-198-846-9854	Sequence 9854, Ap
37	35.4	2.5	2574	13	US-10-311-625-13	Sequence 13, App1
38	35.2	2.5	1115	15	US-10-184-644-440	Sequence 440, App
39	35.2	2.5	1115	15	US-10-184-644-440	Sequence 440, App
40	35.2	2.5	715517	13	US-10-027-632-53712	Sequence 53712, A
41	35.2	2.5	715517	14	US-10-027-632-53712	Sequence 53712, A
42	35.2	2.4	416	11	US-09-918-995-16569	Sequence 16569, A
43	34.6	2.4	1237	13	US-10-027-632-199318	Sequence 199318,
44	34.6	2.4	1237	14	US-10-027-632-199318	Sequence 199318,
45	34.6	2.4	1482	15	US-10-060-036-66	Sequence 66, App1

ALIGNMENTS

RESULT 1

US-09-745-763-35

Sequence 35, Application US/09745763

Patent No. US20020065394A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallee, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Merberg, David
Treacy, Maurice
Spaulding, Vikki

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM

NUMBER OF SEQUENCES: 219

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA

COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/745,763

FILING DATE: 18-Jun-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

```

; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1851 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-745-763-35

Query Match      100.0%; Score 1416; DB 9; Length 1851;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1  ATGAATTCCTTATCTTGCGATTTTCCGTCGTCCTTATACCTTATACCTGTCCTGGG 60
99  ATGAATTCCTTATCTTGCGATTTTCCGTCGTCCTTATACCTTATACCTGTCCTGGG 158
61  AAAGCTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAAATAAAGAAATA 120
159  AAAGCTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAAATAAAGAAATA 218
121  GCCAGCTGTGAGATGTTGCTTAAAGCAATCATCAACTAGCTGTTTATGTAAGCCAG 180
219  GCCAGCTGTGAGATGTTGCTTAAAGCAATCATCAACTAGCTGTTTATGTAAGCCAG 278
181  AACAGATCTTATGAGCGATGTCCTTGTGTTGATCTGTTGAGACCCAGACTGAGTGC 240
279  AACAGATCTTATGAGCGATGTCCTTGTGTTGATCTGTTGAGACCCAGACTGAGTGC 338
241  TCCAGAAACCTTAAAGAAAGCCATCCAAATATATGACCAAAACCTGAGAGATGGGCTG 300
339  TCCAGAAACCTTAAAGAAAGCCATCCAAATATATGACCAAAACCTGAGAGATGGGCTG 398
301  GAGAAAGTTTCACTTGTGAGCGCAATGATATCCCACTGTGGAGAGAGAGAGATCACT 360
399  GAGAAAGTTTCACTTGTGAGCGCAATGATATCCCACTGTGGAGAGAGAGATCACT 458
361  GTGATGCTGAGCGCAAGATTCATTAAGATGATCCTCGGCTGTTGGAGAGAGATGGG 420
459  GTGATGCTGAGCGCAAGATTCATTAAGATGATCCTCGGCTGTTGGAGAGAGATGGG 518
421  ACTCTTCCAGAGAGGATTAACAGCAAGATTCGTGTGTGATCCTCTTGGATGAATGAG 480
519  ACTCTTCCAGAGAGGATTAACAGCAAGATTCGTGTGTGATCCTCTTGGATGAATGAG 578
481  AGAAGGCTCTCAAGAGAGGAGATGTTGTTTATTAACCAACTTATCACTAC 540
579  AGAAGGCTCTCAAGAGAGGAGATGTTGTTTATTAACCAACTTATCACTAC 638
541  TCAAGGAGGTCATATCCGAAACGAGGGGGGTGGAAGCTGCCAAGGTGGGGCTTGG 600
639  TCAAGGAGGTCATATCCGAAACGAGGGGGGTGGAAGCTGCCAAGGTGGGGCTTGG 698
601  GCATCTCATTCATCCGTCGTCCTCTTCCATCTCAAGTCTCACAAGGATTTAG 660
699  GCATCTCATTCATCCGTCGTCCTCTTCCATCTCAAGTCTCACAAGGATTTAG 758
661  GAATACCAAGATGCGTCGCCAAATATCCACAGCCTGTATTACGTTGGAAGATCAGAA 720
759  GAATACCAAGATGCGTCGCCAAATATCCACAGCCTGTATTACGTTGGAAGATCAGAA 818
721  ATGATGTCAAGATGCGTCCTTCTATGAGATCAAAATTTGATTCACCTAAAGTGGGGCA 780
819  ATGATGTCAAGATGCGTCCTTCTATGAGATCAAAATTTGATTCACCTAAAGTGGGGCA 878
781  AAGACCTACCCAGATGATCTTCTTCAACACTGTAGAGAGATCACTGGAGCAATAT 840
879  AAGACCTACCCAGATGATCTTCTTCAACACTGTAGAGAGATCACTGGAGCAATAT 938
841  CCAGAACAGGTTGATCTGTGAGTGAATCTGACAGCTGGAGATGTTGGAGGGTCC 900
939  CCAGAACAGGTTGATCTGTGAGTGAATCTGACAGCTGGAGATGTTGGAGGGTCC 998
```

```

QY 901 ATGATGATGCGCGTGGAGCCCTTATATATGAGAGACACTCTCACTTATTAAGATCTT 960
DB 999 ATGATGATGCGCGTGGAGCCCTTATATATCATGAGAGACCTCACTTATTAAGATCTT 1058
QY 961 GGGCTGCGTCCAAAGAGACTCTGCGGCTGTGCTGTGAGCTGAGAGAAACAAGTGA 1020
DB 1059 GGGCTGCGTCCAAAGAGACTCTGCGGCTGTGCTGTGAGCTGAGAGAAACAAGTGA 1118
QY 1021 GTTGTGCTTCCAGTATATATCACTTACCAAGATTAATTTTCCAACTACAGCTGTGTG 1080
DB 1119 GTTGTGCTTCCAGTATATATCACTTACCAAGATTAATTTTCCAACTACAGCTGTGTG 1178
QY 1081 ATGAGTCTGACGAGAGACCTTCTTACCACTGAGGCTCAATCACTGAGAGAAAG 1140
DB 1179 ATGAGTCTGACGAGAGACCTTCTTACCACTGAGGCTCAATCACTGAGAGAAAG 1238
QY 1141 GCCAGGCGCATCATGAGAGAGTATGAGCCTGTGAGAGCCCTCAATATCATCACTGAGTGC 1200
DB 1239 GCCAGGCGCATCATGAGAGAGTATGAGCCTGTGAGAGCCCTCAATATCATCACTGAGTGC 1298
QY 1201 CTGAGCCATGAGAGAGAGAGAGACATCACTTTTGAATCCAGCTGAGAGTCTGTGAGCC 1260
DB 1299 CTGAGCCATGAGAGAGAGAGAGACATCACTTTTGAATCCAGCTGAGAGTCTGTGAGCC 1358
QY 1261 AGTCTACTGATGATGATATCAAGATTTTCTTCCATCATCTGCCACGAGACACCATG 1320
DB 1359 AGTCTACTGATGATGATATCAAGATTTTCTTCCATCATCTGCCACGAGACACCATG 1418
QY 1321 ACTGTATGATATCCAAAGAGAGATGATGCTGCTGTGTTGGCTGTGTTCTTAT 1380
DB 1419 ACTGTATGATATCCAAAGAGAGATGATGCTGCTGTGTTGGCTGTGTTCTTAT 1478
QY 1381 GTTGTGCAACATGAGAGAAATGCTGCTTACGTTCC 1416
DB 1479 GTTGTGCAACATGAGAGAAATGCTGCTTACGTTCC 1514

RESULT 2
US-09-984-271-28
; Sequence 28, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OR INVENTION: 71 Human Secreted Proteins
;   FILE REFERENCE: P2030P1
;   CURRENT FILING DATE: US/09/984,271
;   PRIOR APPLICATION NUMBER: 09/482,273
;   PRIOR FILING DATE: 2000-01-13
;   PRIOR APPLICATION NUMBER: PCT/US99/15849
;   PRIOR FILING DATE: 1999-07-14
;   PRIOR APPLICATION NUMBER: 60/092,921
;   PRIOR FILING DATE: 1998-07-15
;   PRIOR APPLICATION NUMBER: 60/092,922
;   PRIOR FILING DATE: 1998-07-15
;   PRIOR APPLICATION NUMBER: 60/092,956
;   NUMBER OF SEQ ID NOS: 267
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 28
;   LENGTH: 1863
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-984-271-28

Query Match      99.8%; Score 1412.8; DB 11; Length 1863;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1  ATGAATTCCTTATCTTGCGATTTTCCGTCGTCCTTATACCTTATACCTGTCCTGGG 60
99  ATGAATTCCTTATCTTGCGATTTTCCGTCGTCCTTATACCTTATACCTGTCCTGGG 158
```

QY 61 AAAGCTATATGCAAGATGCGATCTCTAAGAGACCTTTTGAAGAAATTAAGAGAAATA 120
DB 159 AAAGCTATATGCAAGATGCGATCTCTAAGAGACCTTTTGAAGAAATTAAGAGAAATA 218
QY 121 GCCAGCTGTGAGATGTTGCTTAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 180
DB 219 GCCAGCTGTGAGATGTTGCTTAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 278
QY 181 AACGATCTATATGAGCGATGCGCACTTCTGTGTATCTGTGTAAGCCAGCTAGTGGC 240
DB 279 AACGATCTATATGAGCGATGCGCACTTCTGTGTATCTGTGTAAGCCAGCTAGTGGC 338
QY 241 TCCAGAGACCTAAGAAAAAGCCATCCAAATTATGTACCAAACTGACAGCAATGAGGCTG 300
DB 339 TCCAGAGACCTAAGAAAAAGCCATCCAAATTATGTACCAAACTGACAGCAATGAGGCTG 398
QY 301 GAGAAAGTTCACTGTGAGCCAGTGAAGATATCCCACTGGAGAGAGGAGAAATCAAGCT 360
DB 399 GAGAAAGTTCACTGTGAGCCAGTGAAGATATCCCACTGGAGAGAGGAGAAATCAAGCT 458
QY 361 GTGATGCTGAGCCGAAGATTCATAGATAGCCATCTCTGGCTCTTGGAGAGAGATGGG 420
DB 459 GTGATGCTGAGCCGAAGATTCATAGATAGCCATCTCTGGCTCTTGGAGAGAGATGGG 518
QY 421 ACTCTCCAGAGGSCATTTACAGAGAGTTCTGTGTGAGACCTCTTTGATGAATGAG 480
DB 519 ACTCTCCAGAGGSCATTTACAGAGAGTTCTGTGTGAGACCTCTTTGATGAATGAG 578
QY 481 AGAAGGSCCTCAGAGAGCAAGAGAGATGTTGTTTATTAACAACCTTATCATCACTAC 540
DB 579 AGAAGGSCCTCAGAGAGCAAGAGAGATGTTGTTTATTAACAACCTTATCATCACTAC 638
QY 541 TCAAGAGCGGTGCAATATCCGAACGAGGGGCGGTGGAAGCTGCCAAGTGGGGCTTTG 600
DB 639 TCAAGAGCGGTGCAATATCCGAACGAGGGGCGGTGGAAGTGGCCAAGTGGGGCTTTG 698
QY 601 GCATCTCTCATTCGATCCGTGGCTCTCTTCCCATCTAAGTCTCCACAGAGTATTCAG 660
DB 699 GCATCTCTCATTCGATCCGTGGCTCTCTTCCCATCTAAGTCTCCACAGAGTATTCAG 758
QY 661 GAATACAGAGATGCGGTGCCCAAAATTCACAAGCTGTATTAAGTGAAGATGACAGAA 720
DB 759 GAATACAGAGATGCGGTGCCCAAAATTCACAAGCTGTATTAAGTGAAGATGACAGAA 818
QY 721 ATGATGTCAGAAATGCGCTTCTCATGAGATCAAAATTTGTCATTCACTTAAGATGGGCGCA 780
DB 819 ATGATGTCAGAAATGCGCTTCTCATGAGATCAAAATTTGTCATTCACTTAAGATGGGCGCA 878
QY 781 AAGACCTACCCAGATCTGATTCCTTCAACAAGCTGAGAGATCACTGGAGAGAAATAT 840
DB 879 AAGACCTACCCAGATCTGATTCCTTCAACAAGCTGAGAGATCACTGGAGAGAAATAT 938
QY 841 CCAGAAAGAGTTGTACTGTGTCAGTGAATCTGAGACAGCTGGAGTGTGGCAGGGTCC 900
DB 939 CCAGAAAGAGTTGTACTGTGTCAGTGAATCTGAGACAGCTGGAGTGTGGCAGGGTCC 998
QY 901 ATGATGATGCGCGGTGAGCCTTTATATCATGAGAGACCTCTCACTTATTAAGATCTT 960
DB 999 ATGATGATGCGCGGTGAGCCTTTATATCATGAGAGACCTCTCACTTATTAAGATCTT 1058
QY 961 GGGGTGCTCCAAAGAGAGACTCTGCGGTGTGTCTGAGATGAGAGAAAGCAAGTGGGA 1020
DB 1059 GGGGTGCTCCAAAGAGAGACTCTGCGGTGTGTCTGAGATGAGAGAAAGCAAGTGGGA 1118
QY 1021 GTTGTGCTCTTCAGATATATCATGTTACACAAGGTAAATATTTCCACTAGCTGGTG 1080
DB 1119 GTTGTGCTCTTCAGATATATCATGTTACACAAGGTAAATATTTCCACTAGCTGGTG 1178
QY 1081 ATGAGATCTGAGAGAGAACTTTCTTACCACTGAGCTGCAATTCAGTGGAGAAAG 1140
DB 1179 ATGAGATCTGAGAGAGAACTTTCTTACCACTGAGCTGCAATTCAGTGGAGAAAG 1238

QY 1141 GCCAGGCGCATCATGAGAGAGGTTATGAGCTGTGAGGCCCTCAATATCACTAGTTC 1200
DB 1239 GCCAGGCGCATCATGAGAGAGGTTATGAGCTGTGAGGCCCTCAATATCACTAGTTC 1298
QY 1201 CTGAGCCATGAGAGAGGAGACAGATCACTTTTGGATCCAAAGCTGGAGTCTGAGCC 1260
DB 1299 CTGAGCCATGAGAGAGGAGACAGATCACTTTTGGATCCAAAGCTGGAGTCTGAGCC 1358
QY 1261 AGTCTACTGATGATCATATCAAGTATTTCTTCCATCACTACCTCCAGAGACCATG 1320
DB 1359 AGTCTACTGATGATCATATCAAGTATTTCTTCCATCACTACCTCCAGAGACCATG 1418
QY 1321 ACTGTCAATGATCCAAAGAGATGATGTTGCTGCTGCTGTTGGGCTGTTCTTAT 1380
DB 1419 ACTGTCAATGATCCAAAGAGATGATGTTGCTGCTGCTGTTGGGCTGTTCTTAT 1478
QY 1381 GTTGTGCAAGATGAGAAAGATGCTGCTAGTCC 1416
DB 1479 GTTGTGCAAGATGAGAAAGATGCTGCTAGTCC 1514

RESULT 3
US-09-917-800A-505
Sequence 505, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
APPLICANT: Gene Logic, Inc.
TITLE OR INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 505
LENGTH: 1778
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 AF097723
US-09-917-800A-505

Query Match 78.2%; Score 1107.2; DB 10; Length 1778;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 1223; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 1 ATGAATCTCTTATCTTGGATTTTGGTGTGTCACTTTATCTCTGTGCTGGG 60
DB 121 ATGAGGTTCTTTCTTCTGTTGCTGTGTTGTTCACTTTCTCTGGGCTCTGGA 180
QY 61 AAAGCTATATGCAAGATGCGATCTCTAAGAGACCTTTTGAAGAAATTAAGAGAAATA 120

Db 181 AAAGTATATACAAAGTGTGTTCTCAGGCAATTTCAAGAAATAAAGAAATA 240
Qy 121 GCCAGCTGAGAGTGTGCTTAAGACATCATCACTAGCTGTTATGTAAGCCAG 180
Db 241 GCCAATATGAAAGATGTGCTTAAGCAATTAATCACTGCTGTTATGAAATAACAG 300
Qy 181 AACAGATCTTATGAGAGTGTGCACTTCTGTTGATCTGTTGAGCCAGCTAGTGGC 240
Db 301 AACCGTGTATGAGCGCTTGGGACTTCTAGTTGATCTGTTGAGCCAGCTAGTGGC 360
Qy 241 TCCAGAACTTAAAGAAAGCAATCCAAATTAATGACAAAACCTGACAGCAAGTGGCTG 300
Db 361 TCTAAGAACTTAAAGAAAGCTATCCAAATCATGTACAAAACCTGACCAAGATGGCTG 420
Qy 301 GAGAAAGTTCACTGAGAGCAAGTGAATACCCCACTGAGAGAGGAGAAAGATCACT 360
Db 421 GAAAACTGCACTGAGAGCAAGTGAATACCTCACTGAGAGAGGAGGAGAAATCTGCA 480
Qy 361 GTGATGCTGAGAGCAAGTGAATACCTCACTGAGAGAGGAGGAGCACTTGGG 420
Db 481 GTGATGCTGAGAGCAAGTGAATACCTCACTGAGAGAGGAGGAGCACTTGGG 540
Qy 421 ACTCTCTCAGAGAGCAATTAAGCAAGAGTCTGAGTGAACCTCTTTCATGAACTGAG 480
Db 541 ACTCTCTCAGAGAGTATCAAGCAAGAGTCTGAGTGAACCTCTTTCATGAACTGAA 600
Qy 481 AGAAGAGGCTCAGAGAGCAAGAGAGTGTGTTTATTAACCACTTATCATCACTAC 540
Db 601 AGAAGAGGCTCAGAGAGCAAGAGAGTGTGTTTATTAACCACTTATCATCACTAC 660
Qy 541 TCAAGAGGCTCAGAGAGCAAGAGAGTGTGTTTATTAACCACTTATCATCACTAC 600
Db 661 GGGAGAACTGTGCACTACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Qy 601 GCATCTCTCATTCATTCAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 721 GCATCTCTCATTCAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Qy 661 GATATCCAGAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Db 781 GATATCCAGAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Qy 721 ATGATGTCAAGAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 841 ATGATGTCAAGAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Qy 781 AAGACCTTACCAAGATGATTCCTTCAACACTGATGAGAGAGATCACTGGAGCAAAATAT 840
Db 901 AAGACCTTACCAAGATGATTCCTTCAACACTGATGAGAGATCACTGGAGCAAAATAT 960
Qy 841 CCAGAACAGGTTGATCTGCTGAGTGAATCTGCAAGCTGGAGAGTGGGAGGAGGAGG 900
Db 961 CCAGAACAGGTTGATCTGCTGAGTGAATCTGCAAGCTGGAGAGTGGGAGGAGGAGG 1020
Qy 901 ATGATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
Db 1021 CTGATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
Qy 961 GGGCTGCTCCAAAGAGAGCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 1081 GGGCTGCTCCAAAGAGAGCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Qy 1021 GTTGTGCTCTCCAGTATTAAGAGTATTAAGAGTATTAAGAGTATTAAGAGTATTA 1080
Db 1141 GTTGTGCTCTCCAGTATTAAGAGTATTAAGAGTATTAAGAGTATTAAGAGTATTA 1200
Qy 1081 ATGAGTGTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Db 1201 ATGAGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
Qy 1141 GCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200

Db 1261 GCCAGGCTATCATGAAGAGTCAAGTCTCTGCAACCCCTCAATATACCAAGTCT 1320
Qy 1201 CTGAGCATGAGAGAGGAGAGCATCACTTTTGAATCCAGCTGAGTGGCTGAGGCT 1260
Db 1321 TTATATGATGAGAGAGAGTCAATTAATCTTCTGATCCAGCTGAGTGGCTGAGGCT 1380
Qy 1261 AGTCTACTGATGAGAGAGTCAATTAATCTTCTGATCCAGCTGAGTGGCTGAGGCT 1320
Db 1381 AGTCTGAGAGTCAATTAATCTTCTGATCCAGCTGAGTGGCTGAGGCTGAGGCT 1440
Qy 1321 ACTGTCATGATCCAGAGAGTCAATTAATCTTCTGATCCAGCTGAGTGGCTGAGGCT 1380
Db 1441 ACTGTCATGATCCAGAGAGTCAATTAATCTTCTGATCCAGCTGAGTGGCTGAGGCT 1500
Qy 1381 GTTGTGAG 1440
Db 1501 GTTGTGAG 1560

RESULT 4

US-09-984-271-95
; Sequence 95, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 95
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-271-95

Query Match 54.5%; Score 772.4; DB 11; Length 1134;
Best Local Similarity 99.7%; Pred. No. 1.6e-240;
Matches 784; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 631 TCCATCTACAGTCTCTCAACAGGATTCAGAGATCCAGAGTGGCTGCCCAAAATTTCA 690
Db 1 TCCATCTACAGTCTCTCAACAGGATTCAGAGATCCAGAGTGGCTGCCCAAAATTTCA 60
Qy 691 ACAGCCGTATTAAGGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 750
Db 61 ACAGCCGTATTAAGGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 120
Qy 751 AAAATTTCTATTCAGTAAAGATGGGGGCAAGAGCTTACCAATCTGATTCCTTCAAC 810
Db 121 AAAATTTCTATTCAGTAAAGATGGGGGCAAGAGCTTACCAATCTGATTCCTTCAAC 180
Qy 811 ACTGTACAGAGATCACTGGAGAGCAATATTCAGAGAGGTTGATCTGCTGAGTGAAT 870
Db 181 ACTGTACAGAGATCACTGGAGAGCAATATTCAGAGAGGTTGATCTGCTGAGTGAAT 240
Qy 871 CTGAGAGCTGGAGATGTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 930
Db 241 CTGAGAGCTGGAGATGTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
Qy 931 TGGAGACATCTCATCTTATTAAGATCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 990

Db	301	TGGAGAGCACTCTCACTATTAAAGATCTTGGGCTGCGTCCAAAGAGGACCTGTGGCGTG	360
Qy	991	GTGCTCTGAGATCTGAGAAAGAAACAAGTGGAGTTGGTGCTTCCAGTATTATTCAGTTACAC	1056
Db	361	GTGCTCTGAGATCTGAGAAAGAAACAAGTGGAGTTGGTGCTTCCAGTATTATTCAGTTACAC	420
Qy	1051	AAGGTAATAATTTTCCACTACAGTCTGGTGAATGAGCTGACGAGAAACCTTCTTACC	1110
Db	421	AAGGTAATAATTTTCCACTACAGTCTGGTGAATGAGCTGACGAGAAACCTTCTTACC	480
Qy	1111	ACTGGGCTGCATTTCACTGGCACTGTAAGAAAGCCAGGSCCATTCATGAGAGAGTTATGAGC	1170
Db	481	ACTGGGCTGCATTTCACTGGCACTGTAAGAAAGCCAGGSCCATTCATGAGAGAGTTATGAGC	539
Qy	1171	CTGCTGCAGCCCTCAATATCACTCAGGTCCTGAGCCATGAGAGAAAGGACAGACATCAAC	1230
Db	540	CTGCTGCAGCCCTCAATATCACTCAGGTCCTGAGCCATGAGAGAAAGGACAGACATCAAC	599
Qy	1231	TTTTGGATCCAAAGCTGAGAGTGCCTGGAGCCAGTCTACTTGATGACTTATTAACAAGTATTC	1290
Db	600	TTTTGGATCCAAAGCTGAGAGTGCCTGGAGCCAGTCTACTTGATGACTTATTAACAAGTATTC	659
Qy	1291	TTCTTTCATCACTCCACGAGAGACACCATGACCTGTCATGATATCCAAACACATGAATGTT	1350
Db	660	TTCTTTCATCACTCCACGAGAGACACCATGACCTGTCATGATATCCAAACACATGAATGTT	719
Qy	1351	GCTGCTGCTGTTTGGGCTGTTGTTCTTATGTTGTTGTCAGACATGGAAGAAATGCTGCCT	1410
Db	720	GCTGCTGCTGTTTGGGCTGTTGTTCTTATGTTGTTGTCAGACATGGAAGAAATGCTGCCT	779
Qy	1411	AGGTCC 1416	
Db	780	AGGTCC 785	

```

RESULT 5
US-09-833-381-1929
; Sequence 1929, Application US/09833381
; Patent No. US20020132090A1
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
CURRENT FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1929
LENGTH: 357
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(357)
OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1929

```

	Query Match	Similarity	18.5%	Score 261.8	DB 10	Length 357	
	Best Local	Similarity	86.4%	Pred. No. 2.6e-74			
	Matches	298	Conservative	0	Mismatches 46	Indels 1	
					Gaps	14	
QY	970	CCAAAGAGAC	CTCGCGGCTG	GTCTCGAG	CTGCAGAA	GAACAAGATGGATGGTCC	102
Db	13	CGAAAGAGAC	CTCGCGCTGTG	CTCGAG	CTGCAGAA	GAACAAGATGGTCC	72
QY	1030	TTCCAGTATTA	CTAGTTACACAAG	TAATAATTTCC	CAACTAC	AGTCGTGTATGGAGCT	108
Db	73	TCCCAAGTATTA	AGACTACATTAAG	CAATAATTTTCC	CAAGTAC	AGTTGTGTATGGAACT	132
QY	1090	GACGAGAGAC	CTTTTACCTCC	CTGGGTGCA	TTCACTCG	CAGTGAAGGACGAGGCC	114

Db	133	GACTCAGGAACCTTCTTATCCACTGGACTGCACTGCAAGTTTACTGGCAATGACAAAGCCAGGGCT	192
Qy	1150	ATCATGAGAGAGGTTATGAGCGCTGCTGCAAGCCCTCAATATCACTCAGATCCTGAGCAAT	1209
Db	193	ATCATGAAGGAAGTCATGAAATCTTCTGCAACCCCTCAATGTCACCAAGGTCTTTTGTAAAT	252
Qy	1210	GGAAAGGAGACATCAATCTTTTGGATCCAAAGC-TGGAAATGCTGGAGCAGACTACT	1268
Db	253	GGAAAGGAATGATATTAACTTCTTGAAATCCAAAGCTTGGAAATCCTGGAGCCAGATCGCN	312
Qy	1269	TGATGACTTATACAGATTCTTCTTCCATCACTCCACCGAGAGA	1313
Db	313	TGATGACTTGTACAAGATTNCTTTTCCATCACTCCACCGAGAGA	357

```

RESULT 6
US-09-833-381-1930
: Sequence 1930, Application US/09833381
: Patent No. US20020132090A1
: GENERAL INFORMATION:
: APPLICANT: Robison, Keith E.
: TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
: FILE REFERENCE: 5800-119
: CURRENT APPLICATION NUMBER: US/09/833.381
: CURRENT FILING DATE: 2001-04-11
: PRIOR APPLICATION NUMBER: 09/516,448
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 2050
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1930
: LENGTH: 427
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(427)
: OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1930

```

Query Match	17.4%	Score 246.2	DB 10	Length 427
Best Local Similarity	87.3%	Pred. No. 3.5e-69		
Matches 269	Conservative 0	Mismatches 39	Indels 0	Gaps 0
Qy	33	TGTTACCTTTATCCCTGTGCTCTGTGGAAAGCTATATGCAAGATGGCATCTCTAAG	92	
Db	117	TGTTACCTTTTACCTTGGGCTCTGGTAAAGCTGTATTCAGATGGTGTCTCAGCG	176	
Qy	93	GACTTTGGAGAAATATAAGAAATATGCCAGCTGTGGAGATGTTGTTAAGCAATCAT	152	
Db	177	AACATTTTCGAAATATAAGAAAGAAATAGCCATATGAAAGATGTGTGTAAGCAATTAT	236	
Qy	153	CAACCTAGCTGTTATGTAAGTAAAGCCCAAGACAGATCCATGAGGATGGCACTTCGT	212	
Db	237	CAACCTGCTGTTTATGTATGTAATATCCAGAACCGGTCTATGAGCGCTTTGGGACTTCTAGT	296	
Qy	213	TGATACTGTTGAGCCAGACTGAGTGGCTCCAGAACCTAGAAAAAGCATCCAATTAT	272	
Db	297	TGATAGCTGTTGAGCCAGACTGAGTGGCTCTAAGAACCTAGAAAAAGCATTTCAATCAT	356	
Qy	273	GTACCAAAACCTGAGCAAGATGGGCTGTGAGAAAGTTCACTTGGAGCCAGTGAAGATACC	332	
Db	357	GTACCAAAACCTGAGCAAGATGGGCTTAGAAAAATTTCACTTGGAGCGAGTCANAAATACC	416	
Qy	333	CCACTGGG 340		
Db	417	CCACTGGG 424		

RESULT 7
 US-09-783-590-12141
 ; Sequence 12141, Application US/09783590
 ; Patent No. US20020110850A1
 ; GENERAL INFORMATION:

```

; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16, 2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12141
; LENGTH: 134
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (29)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (114)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (118)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (126)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-12141

Query Match      6.4%; Score 91; DB 10; Length 134;
Best Local Similarity 83.5%; Pred. No. 6.9e-19;
Matches 111; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

Oy      535  AACTACTAAGAGACGTCACATACCGAAGCAGAGGCGGTGAGAGCTGCCAAGGTGGG 594
      |||
Db      2  ANCTACTAAGAGACGTCACATACCGAAGCAGAGGCGGTGAGAGCTGCCAAGGTGGG 61

Oy      595  GCTTTGGCATCTCTCATTCG-ATCCGTGGCTCTTTCATCTACAGTCTCACACAG 653
      |||
Db      62  GCTTTGGCATCTCTTATTTAGGATCGGTGGCTCTTTCATCTAAGCTCTMAANAG 121

Oy      654  TATTCAGGAATAC 666
      |||
Db      122  TTTNAGGATAC 134

RESULT 8
US-09-908-975-25931
; Sequence 25931, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSEMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC
; FILE REFERENCE: 36688-0005

```

```

; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25931
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-908-975-25931

Query Match      4.3%; Score 60.2; DB 13; Length 65;
Best Local Similarity 95.4%; Pred. No. 4.8e-09;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      1333  CCAAGCAGATGATGTTGCTGCTGTTGGGCTGTTCTTATGTTGTCAGAC 1392
      |||
Db      1  CCAAGCAGATGATGTTGCTGCTGTTGGGCTGTTCTTATGTTGTCAGAC 60

Oy      1393  ATGGA 1397
      |||
Db      61  ATGGA 65

RESULT 9
US-09-908-975-8960
; Sequence 8960, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSEMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8960
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-908-975-8960

Query Match      4.2%; Score 60; DB 13; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      152  TCAACTAGCTGTTTATGTTAAGCCAGAACAGATCTTATGACGATGGCACTTCTG 211
      |||
Db      1  TCAACTAGCTGTTTATGTTAAGCCAGAACAGATCTTATGACGATGGCACTTCTG 60

RESULT 10
US-10-059-585-5
; Sequence 5, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru

```

APPLICANT: Yamamoto, Jun-ichi
APPLICANT: Ishii, Shizuko
APPLICANT: Sugiyama, Tomoyasu
APPLICANT: Makamatsu, Ai
APPLICANT: Nagai, Keiichi
APPLICANT: Otsuki, Tetsuji
APPLICANT: Funahashi, Shin-ichi
APPLICANT: Senoo, Chiaki
APPLICANT: Nezu, Jun-ichi
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
FILE REFERENCE: 06501-098001
CURRENT APPLICATION NUMBER: US/10/059,585
CURRENT FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: PCT/JP00/05060
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/183,322
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: US 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 11-248036
PRIOR FILING DATE: 1999-07-29
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (173)...(1450)
US-10-059-585-5

Query Match 3.0%; Score 42; DB 15; Length 2200;
Best Local Similarity 45.5%; Pred. No. 0.04;
Matches 150; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 835 AATATATCAGAAAGCTTGTATCTGTCAGTGCAGATCTGCAGACCTGGATGTTGGCCAG 894
DB 158 AAGGATGAGAGGAGGATGTCAGAGGAGAGAGCTGCTGATGTTTCATGGGGGCTGAGGAG 217
QY 895 GGTGTCATGATGATGATGCGCGTGGAGCTTTATATCATGGAAGCACTCTCATTTATAA 954
DB 218 GCAGCCCTGACCCAGCCGAGCGGAGTGGCGGAGAGGGGGTGGCTGGGCTCTTTCAGGG 277
QY 955 GATCTTGGGCTGCGTCCAAAGAGACTCTGCGGCTGGTCTGGAATGCAAGAAACAA 1014
DB 278 GAGAGAGGCGACGCTCTCATGTCAGAAAGTGTGCTGCTGCTTCAAGTGAAGAGAGGA 337
QY 1015 GGTGAGATTGTGCTCTTCAGTATTATCATGTTAAATATTTCCAACTACAGT 1074
DB 338 GGAAGAGAGAGTCCCTGAGATTTTGTATACCAAGGCTCTCGGCCCGCATCAGC 397
QY 1075 CTGATGATGAGTGTGACGCGCAACTCTTCAACCACTGGGCTGCAATTACTGGCAGT 1134
DB 398 ATCCCTGCTCTTTCATCATCAGACGTCGGAACAACCAAGCCCTGAGATCCTGACCGG 457
QY 1135 GAAGAAGCGAGGCGCATCATGAGAGGTT 1164
DB 458 GAGAACACGTTTGTGTGAAGTGAAGT 487

RESULT 11
US-09-908-975-1202
Sequence 1202, Application US/09908975
Publication No. US20030165843A1
GENERAL INFORMATION:
APPLICANT: SHOSHAN, Avi
APPLICANT: WASSERMAN, Alon

APPLICANT: MINTZ, Eli
APPLICANT: MINTZ, Ilat
APPLICANT: FAIGLER, Simon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
FILE REFERENCE: 36688-0005
CURRENT APPLICATION NUMBER: US/09/908,975
CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1202
LENGTH: 65
TYPE: DNA
ORGANISM: Rattus norvegicus
US-09-908-975-1202

Query Match 2.9%; Score 41; DB 13; Length 65;
Best Local Similarity 76.9%; Pred. No. 0.0087;
Matches 50; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 343 AGGGGAGAAAGATCAGCTGTGATGTCGAGCCCAAGATTCATAGATGATGCTGGT 402
DB 1 AGGGGCAAGAAATCTGACGATGATGCTGCTCGAATTCACAAATGCTATTTAGGC 60
QY 403 CTTGG 407
DB 61 CTTGG 65

RESULT 12
US-09-918-995-30889
Sequence 30889, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30889
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(484)
OTHER INFORMATION: n = A, T, C or G
US-09-918-995-30889

Query Match 2.8%; Score 39.4; DB 11; Length 484;
Best Local Similarity 49.3%; Pred. No. 0.11;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 216 TACTGTGACCCAGATGATGCTTCAGAACTTGAAAGCCATCCAAATTATGTA 275
DB 231 TAGGTCTGTTCCAAAGCACTGTGTACAGAAAGCTGAATTCATTAGAGCAACAAC 290
QY 276 CCAAAACCTGACAGATGAGGCTGAGAAAGTTCACTGAGCACTGAGAAATACCCA 335
DB 291 CCTGAGAAATACACAGAGGCGACGCTTCAGATGATGTGTGGAGAGAGAGGAG 350
QY 336 CTGGAGAGGAGGAGAAATCAGCTGTGATGCTGAGCCCAAGATTCATAGATGACCAT 395
DB 351 AGGGGACAGGGGACAGATTCAGCTTGTGTGGGTCTGAGGGGTCTACAGGGGTAG 410

Oy 396 CTGGGCTCTGGCAGCAGCATTTGGGACTC 424
Db 411 CCAGGATCTGGGAAACAGATCCAGCACTC 439

RESULT 13
US-10-232-484-5
Sequence 5, Application US/10232484
Publication No. US20030113847A1

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Yonghong
APPLICANT: Drmanac, Radote T
TITLE OF INVENTION: No. US20030113847A1el Nucleic Acids and Polypeptides
FILE REFERENCE: 791C1P2BDIV
CURRENT APPLICATION NUMBER: US/10/232,484
CURRENT FILING DATE: 2002-08-30
PRIOR APPLICATION NUMBER: 09/695,783
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: 09/552,929
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1363
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (848)..(1090)
OTHER INFORMATION:
US-10-232-484-5

Query Match 2.8%; Score 39.4; DB 15; Length 1363;
Best Local Similarity 49.3%; Pred. No. 0.2;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Oy 216 TACTGTGGACCCAGCTAGTGGCTCCAGAACTAGAAAAGCCATCCAAATTATGTA 275
Db 877 TAGTGTCTGTTCCAAAGCAGCTGTGTACAGAACTAGAAATTAACCATTAAGGCCAAAC 936
Oy 276 CCAAAACCTGACAGATGGGCTGAGAAAGTTCACTGAGCCAGTGAATACCCCA 335
Db 937 CCTGAGAAATACAAAGGGGGAGCGCTTCAGTAATGTGTTGGGAAAGAGAGCGGAG 996
Oy 336 CTGGAGAGGGGAGAAATCACTGTGATGCTGAGCCAGAAATTCATTAAGATGCCAT 395
Db 997 AGGGGACAGGGGACAGATTCACTTTGTGTGGGTCTCTGAGGGGTTCTTACCAGGGGTAG 1056
Oy 396 CCTGGGCTCTGGCAGCAGCATTTGGGACTC 424
Db 1057 CCAGGATCTGGGAAACAGATCCAGCACTC 1085

RESULT 14
US-10-044-090-579
Sequence 579, Application US/10044090
Publication No. US20020137081A1

GENERAL INFORMATION:
APPLICANT: Olga Bandman
TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
FILE REFERENCE: PA-0028 US
CURRENT APPLICATION NUMBER: US/10/044,090
CURRENT FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 850
SOFTWARE: PERL Program
SEQ ID NO 579
LENGTH: 2319
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. US20020137081A1 023244.1
NAME/KEY: unsure
LOCATION: 169, 173
OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-579

Query Match 2.8%; Score 39.4; DB 14; Length 2319;
Best Local Similarity 49.3%; Pred. No. 0.29;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Oy 216 TACTGTGGACCCAGCTAGTGGCTCCAGAACTTGAAGAAACCCATCCAAATTATGTA 275
Db 1560 TAGTGTCTGTTCCAAAGCAGCTGTGTACAGAACTAGAAATTAACCATTAAGGCCAAAC 1619
Oy 276 CCAAAACCTGACAGATGGGCTGAGAAAGTTCACTGAGCCAGTGAATACCCCA 335
Db 1620 CCTGAGAAATACAAAGGGGGAGCGCTTCAGTAATGTGTTGGGAAAGAGAGCGGAG 1679
Oy 336 CTGGAGAGGGGAGAAATCACTGTGATGCTGAGCCAGAAATTCATTAAGATGCCAT 395
Db 1680 AGGGGACAGGGGACAGGATTCAGCTTTGTGTGGGTCTCTAAGGTTCTTACCAGGGGTAG 1739
Oy 396 CCTGGGCTCTGGCAGCAGCATTTGGGACTC 424
Db 1740 CCAGGATCTGGGAAACAGATCCAGCACTC 1768

RESULT 15
US-09-814-353-20277
Sequence 20277, Application US/09814353
Publication No. US20030165831A1

GENERAL INFORMATION:
APPLICANT: Thompson, Pamela
APPLICANT: Lee, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER
FILE REFERENCE: MRI-006B
CURRENT APPLICATION NUMBER: US/09/814,353
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,820
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 20277
LENGTH: 3322
TYPE: DNA
ORGANISM: Homo sapiens
US-09-814-353-20277

Query Match 2.8%; Score 39.4; DB 13; Length 3322;
Best Local Similarity 49.3%; Pred. No. 0.36;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Oy 216 TACTGTGGACCCAGCTAGTGGCTCCAGAACTTGAAGAAACCCATCCAAATTATGTA 275
Db 1946 TAGTGTCTGTTCCAAAGCAGCTGTGTACAGAACTAGAAATTAACCATTAAGGCCAAAC 2005
Oy 276 CCAAAACCTGACAGATGGGCTGAGAAAGTTCACTGAGCCAGTGAATACCCCA 335
Db 2006 CCTGAGAAATACAAAGGGGGAGCGCTTCAGTAATGTGTGGGAAAGAGAGGGGAG 2065

Qy	336	CTGGGAGAGGGGAGAGATCAGCTGTGATGCTGGAGCCAAAGATTGATAGATAGCCAT	395
Db	2066	AGGGGACAGGGGACAGGATTCAGCTTGTGTGGTCTGAGGGGTCTTAACAAGGGGTAG	2125
Qy	396	CCTGGGTCTTGGGACAGCATTTGGGACTC	424
Db	2126	CCAGGATCTGGGAAACAGATCAGCGACTC	2154

Search completed: December 22, 2003, 23:44:40
Job time : 447.263 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

.OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:18:08 ; Search time 4851.69 Seconds
(without alignments)
11332.658 Million cell updates/sec

Title: US-09-745-763-35_COPY_171_1514
Perfect score: 1344
Sequence: 1 AAGAAATGCATCTCTAAGAG.....AAGAAATGCCTAGGTC 1344

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:.*
1: gb Da:.*
2: gb Htg:.*
3: gb In:.*
4: gb Om:.*
5: gb Ov:.*
6: gb Pat:.*
7: gb Ph:.*
8: gb Pl:.*
9: gb Pr:.*
10: gb Ro:.*
11: gb Sts:.*
12: gb Sy:.*
13: gb Un:.*
14: gb Vi:.*
15: em Ba:.*
16: em Fun:.*
17: em Hum:.*
18: em In:.*
19: em Mu:.*
20: em Om:.*
21: em Or:.*
22: em Ov:.*
23: em Pat:.*
24: em Ph:.*
25: em Pl:.*
26: em Ro:.*
27: em Sts:.*
28: em Un:.*
29: em Vi:.*
30: em Htg Hum:.*
31: em Htg Inv:.*
32: em Htg Other:.*
33: em Htg Mus:.*
34: em Htg Pln:.*
35: em Htg Rod:.*
36: em Htg Mam:.*
37: em Htg Vrt:.*
38: em Sv:.*
39: em Htgo Hum:.*
40: em Htgo Mus:.*
41: em Htgo Other:.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	1851	6	BD106411
2	1340.8	99.8	1778	6	BD105816
3	1340.8	99.8	1860	6	AX006440
4	1340.8	99.8	1928	9	BC020689
5	1337.6	99.5	1472	9	AF107834
6	1336	99.4	1860	6	BD127520
7	1336	99.4	1860	6	AK075132
8	1327.2	98.8	1794	9	AF119386
9	1086.4	80.7	1596	10	BC037067
10	1084.8	80.7	1596	10	AF107835
11	1060.8	78.9	1726	10	AF131077
12	1060.8	78.9	1778	6	AX400829
13	1060.8	78.9	1778	10	AF097723
14	1000.4	74.4	1716	10	AF009513
15	448.6	33.4	609	6	BD059610
16	421.4	31.4	423	9	AF107833
17	398.2	29.6	642	6	BD125210
18	398.2	29.6	642	6	BD126476
19	354.4	27.1	156702	2	AC084223
20	354.4	27.1	157927	9	AP006278
21	364.4	27.1	171301	9	AC010859
22	331	24.6	417	10	AF107832
23	292.8	21.8	214287	2	AC129951
24	279.4	20.8	167883	2	AC121026
25	279.4	20.8	218874	2	AC110103
26	212.8	15.8	314	6	BD058492
27	212.2	15.8	144703	9	AP003112
28	205	15.3	152176	9	AP003117
29	205	15.3	171035	9	AP003111
30	204.6	15.2	462	6	BD105811
31	204.6	15.2	462	6	E62849
32	174.4	13.0	214287	2	AC129951
33	172.8	12.9	203469	2	AC123654
34	171.2	12.7	218874	2	AC110103
35	163.4	12.2	166050	2	AC013817
36	163.4	12.2	177008	2	AC011175
37	163.4	12.2	240178	2	AC094173
38	163.4	12.2	288136	2	AC120489
39	154.4	11.5	238034	2	AC123681
40	140.6	10.5	317	6	BD076934
41	116	8.6	10792	1	AE015790
42	115	8.6	110000	2	AL390202_01
43	114.2	8.5	492	6	BD125961
44	99.2	7.4	10487	1	AE005922
45	96.4	7.2	1729	3	AF077194

ALIGNMENTS

RESULT 1
LOCUS BD106411 1851 bp DNA linear PAT 18-SEP-2002
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD106411
VERSION BD106411.1 GI:23201229
KEYWORDS JP 2002503955-A/2.
SOURCE Chlamydia sp.
ORGANISM Chlamydia sp.
REFERENCE 1 (bases 1 to 1851)
JACOBS,K., MCCOY,J.M., LAVALLIE,E.R., RACIE,L.A., MERBERG,D.,
TREACY,M., SPAULDING,V. and AGOSTINO,M.J.
Secreted proteins and polynucleotides encoding them
Patent: JP 2002503955-A 2 05-FEB-2002;
JOURNAL

COMMENT

GENETICS INSTITUTE INC
 PN JP 2002503955-A/2
 PD 05-FEB-2002
 PF 20-MAR-1998 JP 1998545874
 PR 21-MAR-1997 US 08/822167,19-MAR-1998 US 09/044466 PI
 KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
 DAVID MERBERG,
 PI MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
 C12N15/12,C07K14/47,A61K38/17
 CC Strandness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers.

FEATURES

1..1851
 /organism="Chlamydia sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:35827"

BASE COUNT 531 a 413 c 438 g 469 t
 ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 1851;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAGAATGCGATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 60
 171 AAGAATGCGATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 230
 61 GATGTTGCTAAGCAATCATCACTAGCTGTTATGTTAAAGCCAGAAACAGATCTAT 120
 231 GATGTTGCTAAGCAATCATCACTAGCTGTTATGTTAAAGCCAGAAACAGATCTAT 290
 121 GAGGATTTGGAAGCTTCTGTTGATCTGTGGAAGCCAGATGAGTGTCCAGAAAGCTTA 180
 291 GAGGATTTGGAAGCTTCTGTTGATCTGTGGAAGCCAGATGAGTGTCCAGAAAGCTTA 350
 181 GAAAGAGCCATCCAAATATATGATCAAAACCTGAGCAAGATGGGCTGGAAGAGTTGAC 240
 351 GAAAGAGCCATCCAAATATATGATCAAAACCTGAGCAAGATGGGCTGGAAGAGTTGAC 410
 241 CTGAGAGCAGTGAAGATACCCCACTGGAGAGAGGAGAAAGATCAGCTGTGATGAG 300
 411 CTGAGAGCAGTGAAGATACCCCACTGGAGAGAGGAGAAAGATCAGCTGTGATGAG 470
 301 CCAAGAAATTCATTAAGATAGCCATCTGAGTCTTGGCAGCAGATGGGAGCTTCTTCAAA 360
 471 CCAAGAAATTCATTAAGATAGCCATCTGAGTCTTGGCAGCAGATGGGAGCTTCTTCAAA 530
 361 GGCATTACAGCAGAAAGTTCTGGTGTGACCTCTTCATGAACTGCAAGAAAGGCTTCA 420
 531 GGCATTACAGCAGAAAGTTCTGGTGTGACCTCTTCATGAACTGCAAGAAAGGCTTCA 590
 421 GAAGCAAGAGGAGAAATGTTGTTTAAACCAACCTTACATCACTCAAGAGCGGTG 480
 591 GAAGCAAGAGGAGAAATGTTGTTTAAACCAACCTTACATCACTCAAGAGCGGTG 650
 481 CAATACCCAGAGCGAGGCGGTGAGAGCTGCAAGAGTGGGGCTTTGGCATCTTCAAT 540
 651 CAATACCCAGAGCGAGGCGGTGAGAGCTGCAAGAGTGGGGCTTTGGCATCTTCAAT 710
 541 CGATCCGAGGCTCTTCTCATCTACAGTCTCTCACAGAGATTCAGAAATACAGAGAT 600
 711 CGATCCGAGGCTCTTCTCATCTACAGTCTCTCACAGAGATTCAGAAATACAGAGAT 770
 601 GGCAGTCCCAAAATTCACACAGCTGTATTTACGTTGAAGATGCAAGAAATGATCAAGA 660
 771 GGCAGTCCCAAAATTCACACAGCTGTATTTACGTTGAAGATGCAAGAAATGATCAAGA 830
 661 ATGGCTTCTCATGGAGATCAAAATTTGTCTATTCAGCTTAAAGATGGGGCAAAAGCTTACCA 720
 831 ATGGCTTCTCATGGAGATCAAAATTTGTCTATTCAGCTTAAAGATGGGGCAAAAGCTTACCA 890
 721 GATACCTGATTCCTTCAACACTGTACAGAGATCACTGGGAGAGCAATATTCAGAAACAGTT 780

891 GATACCTGATTCCTTCAACACTGTACAGAGATCACTGGGAGCAAAATATCCAGAAACAGTT 950
 781 GATACCTGATTCCTTCAACACTGTACAGAGATCACTGGGAGCAAAATATCCAGAAACAGTT 840
 951 GATACCTGATTCCTTCAACACTGTACAGAGATCACTGGGAGCAAAATATCCAGAAACAGTT 1010
 841 GATACCTGATTCCTTCAACACTGTACAGAGATCACTGGGAGCAAAATATCCAGAAACAGTT 900
 1011 GATACCTGATTCCTTCAACACTGTACAGAGATCACTGGGAGCAAAATATCCAGAAACAGTT 1070
 901 AAGAGAGCTGCGGCTGTGCTCTGAGCTGCAAGAGAAACAGAGTGGTGTGCTTC 960
 1071 AAGAGAGCTGCGGCTGTGCTCTGAGCTGCAAGAGAAACAGAGTGGTGTGCTTC 1130
 961 CAGTATATCAGTTACCAAGAGTAAATTTTCCACATCAGTCTGTGATGAGTCTGAC 1020
 1131 CAGTATATCAGTTACCAAGAGTAAATTTTCCACATCAGTCTGTGATGAGTCTGAC 1190
 1021 GCAGAGACCTTCTTACCCAGCTGCTGCAATTCATCTGAGTGAAGAGGCAAGGCGCATC 1080
 1191 GCAGAGACCTTCTTACCCAGCTGCTGCAATTCATCTGAGTGAAGAGGCAAGGCGCATC 1250
 1081 ATGAGAGAGGTTATGAGCTGTGCAAGCTGCAAGCTGCAATTCATCTGAGTGAAGAGG 1140
 1251 ATGAGAGAGGTTATGAGCTGTGCAAGCTGCAAGCTGCAATTCATCTGAGTGAAGAGG 1310
 1141 GAAGGAGCAGACATCAACTTTTGGATCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 1200
 1311 GAAGGAGCAGACATCAACTTTTGGATCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 1370
 1201 GACTTATACAGATTTCTTCTTCATCACTCCACAGAGACATCAGTCTGATGAGAT 1260
 1371 GACTTATACAGATTTCTTCTTCATCACTCCACAGAGACATCAGTCTGATGAGAT 1430
 1261 CCAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAATGTTTGCAGAC 1320
 1431 CCAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAATGTTTGCAGAC 1490
 1321 ATGGAAGAAATGCTGCTAGGTC 1344
 1491 ATGGAAGAAATGCTGCTAGGTC 1514

RESULT 2

BD105816

LOCUS

Secretory protein.

BD105816

ACCESSION BD105816.1 GI:23200634

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JACOBS, K. McCoy, J.M., Racie, L.A., Lavallie, E.R., Merberg, D. and

Spaulding, V.

TITLE

JOURNAL

COMMENT

PATENT: JP 2002502234-A 7 22-JAN-2002;

GENETICS INSTITUTE INC

PN JP 2002502234-A/7

PD 22-JAN-2002

PF 16-APR-1997 JP 1997537384

PR 18-APR-1996 US 08/634325,13-JAN-1997 US 08/783520 PI

KENNETH JACOBS,JOHN M MCCOY,LISA A RACIE,EDWARD R LAVALLIE, PI

DAVID MERBERG,

PI VIKKI SPAULDING

PC C12N15/12,C07K14/47,A61K38/17

CC Strandness: Double;

CC Topology: Linear;

FH Key Location/Qualifiers.

FEATURES

source 1..1778

/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

BASE COUNT 514 a 386 c 422 g 456 t

Query Match 99.8%; Score 1340.8; DB 6; Length 1778;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AAGAAATGGCAATCTCTAAGAGACCTTTGAAAGAAATTAAGAAATAGCCAGCTGGA 60
89 AAGAAATGGCAATCTCTAAGAGACCTTTGAAAGAAATTAAGAAATAGCCAGCTGGA 148
61 GATGTTCTAAGAGACATCAACCTAGCTGTTATGTTAAGCCAGAAACAGATCTAT 120
149 GATGTTCTAAGAGACATCAACCTAGCTGTTATGTTAAGCCAGAAACAGATCTAT 208
121 GAGCGATTGGCACTTCTGTTGATGATCTGTTGACCCAGACTGAGTGGCTCCAGAACCTA 180
209 GAGCGATTGGCACTTCTGTTGATGATCTGTTGACCCAGACTGAGTGGCTCCAGAACCTA 268
181 GAAAAAGCCATCCAAATTAATGATCAAAACCTGACGAGATGGGCTGGAGAAAGTTGAC 240
269 GAAAAAGCCATCCAAATTAATGATCAAAACCTGACGAGATGGGCTGGAGAAAGTTGAC 328
241 CTGAGGCACTGAGAAATACCCCACTGGAGAGAGGGGAGAAAGATGAGCTGATGCTGAG 300
329 CTGAGGCACTGAGAAATACCCCACTGGAGAGAGGGGAGAAAGATGAGCTGATGCTGAG 388
301 CCAAGATTCATTAAGATAGCCATCTGAGCTTGGGACAGATGGGACCTCTCCAGAA 360
389 CCAAGATTCATTAAGATAGCCATCTGAGCTTGGGACAGATGGGACCTCTCCAGAA 448
361 GGCATTACAGCAAGATTTCTGTTGATGATCTTTCATGATGACGAGAGAGGCTTCA 420
449 GGCATTACAGCAAGATTTCTGTTGATGATCTTTCATGATGACGAGAGAGGCTTCA 508
421 GAAGCAAGAGGAGATTTGTTTATTAACCACTTAACATCTCAAGAGAGGCTG 480
509 GAAGCAAGAGGAGATTTGTTTATTAACCACTTAACATCTCAAGAGAGGCTG 568
481 CAATACCGAAGCGAGGGGCGGTGGAAGCTGCGCAAGGTGGGGCTTTGGCATCTCTCAT 540
569 CAATACCGAAGCGAGGGGCGGTGGAAGCTGCGCAAGGTGGGGCTTTGGCATCTCTCAT 628
541 CGATCCGTGGCTCTCTCTCTCATCTCAAGCTTCAACAGGTATTCAGAAATACAGGAT 600
629 CGATCCGTGGCTCTCTCTCTCATCTCAAGCTTCAACAGGTATTCAGAAATACAGGAT 688
601 GGGGTGCGCCAAAATTCACAGCGCTGTTATTAACCGTGGAGATTCAGAAATATGTCAGA 660
689 GGGGTGCGCCAAAATTCACAGCGCTGTTATTAACCGTGGAGATTCAGAAATATGTCAGA 748
661 ATGGCTTCTCATGGGATCAAAATTTGATTCAGCTAAAGATGGGGCAAAAGACCTACCA 720
749 ATGGCTTCTCATGGGATCAAAATTTGATTCAGCTAAAGATGGGGCAAAAGACCTACCA 808
721 GATACGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGTT 780
809 GATACGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGTT 868
781 GATACGATTTCTTCAACACTGTAGCAGAGATTCAGCTGGAGCGGTGCTATGATGATGAGC 840
869 GATACGATTTCTTCAACACTGTAGCAGAGATTCAGCTGGAGCGGTGCTATGATGATGAGC 928
841 GGTGAGACCTTTATATCATGGAAGCACTCTCACTTATTAAGAATCTTGGGCTGCTTCA 900
929 GGTGAGACCTTTATATCATGGAAGCACTCTCACTTATTAAGAATCTTGGGCTGCTTCA 988
901 AAGAGACTCTGCGGCTGCTGCTGCACTGCAAGAGAAACAAGGTGAGATGCTTCC 960
989 AAGAGACTCTGCGGCTGCTGCTGCACTGCAAGAGAAACAAGGTGAGATGCTTCC 1048

961 CAGTATTATCAGTTACACAGGTAATATTTCACATACAGTCTGTATGATGATGAC 1020
1049 CAGTATTATCAGTTACACAGGTAATATTTCACATACAGTCTGTATGATGATGAC 1108
1021 GCAGAAACCTTTTACCACTGAGGCTGCAATTCATCTGACGATGAAAGGCCAGGCCATC 1080
1109 GCAGAAACCTTTTACCACTGAGGCTGCAATTCATCTGACGATGAAAGGCCAGGCCATC 1168
1081 ATGAGAGGTTATGAGGCTGAGGCTGCAAGGCTTCAATTCATCTGAGCTGAGCAGTGA 1140
1169 ATGAGAGGTTATGAGGCTGAGGCTGCAAGGCTTCAATTCATCTGAGCTGAGCAGTGA 1228
1141 GAAGGACACATCAACTTTTGAATCCAGTGAAGTGGCTGAGGCTGAGCTTATCTTAT 1200
1229 GAAGGACACATCAACTTTTGAATCCAGTGAAGTGGCTGAGGCTGAGCTTATCTTAT 1288
1201 GACTTATACAAATTTTCTTTCATCACTCCACGAGACACCATGACTGTCATGAT 1260
1289 GACTTATACAAATTTTCTTTCATCACTCCACGAGACACCATGACTGTCATGAT 1348
1261 CCAAGCAATGATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
1349 CCAAGCAATGATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1408
1321 ATGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1344
1409 ATGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1432

RESULT 3
AX006440 1860 bp DNA linear PAT 06-SEP-2000
LOCUS
DEFINITION
Sequence 5 from Patent WO0004157.
AX006440
VERSION
AX006440.1 GI:9994575
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
Neefs,J.M., Peeters,D.C. and Pangalos,M.
Cloning and characterisation of novel mammalian peptidases
JOURNAL
Patent: WO 0004157-A 5 27-JAN-2000;
JANSSEN PHARMACEUTICA NV (BE); NEERES JEAN MARC EDMOND PERNAND (BE);
PEETERS DANIELLE CELINE GEORGE (BE); PANGALOS MENLAS (GB)
location/Qualifiers
1..1860
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 518 a 420 c 451 g 471 t

ORIGIN

Query Match 99.8%; Score 1340.8; DB 6; Length 1860;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AAGAAATGGCAATCTCTAAGAGACCTTTGAAAGAAATTAAGAAATAGCCAGCTGGA 60
197 AAGAAATGGCAATCTCTAAGAGACCTTTGAAAGAAATTAAGAAATAGCCAGCTGGA 256
61 GATGTTCTAAGAGACATCAACCTAGCTGTTATGTTAAGCCAGAAACAGATCTAT 120
257 GATGTTCTAAGAGACATCAACCTAGCTGTTATGTTAAGCCAGAAACAGATCTAT 316
121 GAGCGATTGGCACTTCTGTTGATGATCTGTTGACCCAGACTGAGTGGCTCCAGAACCTA 180
317 GAGCGATTGGCACTTCTGTTGATGATCTGTTGACCCAGACTGAGTGGCTCCAGAACCTA 376
181 GAAAAAGCCATCCAAATTAATGATCAAAACCTGACGAGATGGGCTGAGAAAGTTGAC 240

```

Db      377  GAAAAAGCCATCAAAATTATGATCCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCAAC 436
Qy      241  CTGAGCCAGTGAAGTAATCCCACTGGAGAGAGGAGAAAGTAATCAAGCTGTGATGCTGGAG 300
Db      437  CTGAGGACGAGAGAAATACCCCACTGGAGAGAGGAGAAAGTAATCAAGCTGTGATGCTGGAG 496
Qy      301  CCAAGAATTCAAAAGATAGCCATCTGGGCTTGGCAGACAGATGGGAGCTCTCCACAA 360
Db      497  CCAAGAATTCAAAAGATAGCCATCTGGGCTTGGCAGACAGATGGGAGCTCTCCACAA 556
Qy      361  GGCATTACAGCAAGAGTTCTGGTGTGAACCTTTTGATGAATCTGCAGAGAGAGGCTCA 420
Db      557  GGCATTACAGCAAGAGTTCTGGTGTGAACCTTTTGATGAATCTGCAGAGAGAGGCTCA 616
Qy      421  GAAGCAAGAGGAAAGTTGTTTATATACCACTTACATCACTACTCAAGAGAGGCTG 480
Db      617  GAAGCAAGAGGAAAGTTGTTTATATACCACTTACATCACTACTCAAGAGAGGCTG 676
Qy      481  CAATACCGAAGCAGAGGGGCGGTGGAGCTGCAGAGTGGGGGCTTTGGCATCTCTCAT 540
Db      677  CAATACCGAAGCAGAGGGGCGGTGGAGCTGCAGAGTGGGGGCTTTGGCATCTCTCAT 736
Qy      541  CGATCCGTCGCTCTCTCTCTCACTACAGTCTCAACAGGATTCAGAGAAATACAGAGAT 600
Db      737  CGATCCGTCGCTCTCTCTCTCACTACAGTCTCAACAGGATTCAGAGAAATACAGAGAT 796
Qy      601  GCGCGTCCCAAAATTCACAAGCCTGTATTAACGGTGGAAAGTGCAGAAATATGTCAGA 660
Db      797  GCGCGTCCCAAAATTCACAAGCCTGTATTAACGGTGGAAAGTGCAGAAATATGTCAGA 856
Qy      661  ATGGCTTCTCAATGGGATCAAAATTTGTCATTCAGTCAAGTGGGGGCAAGACCTACCCA 720
Db      857  ATGGCTTCTCAATGGGATCAAAATTTGTCATTCAGTCAAGTGGGGGCAAGACCTACCCA 916
Qy      721  GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCAAGACAGTT 780
Db      917  GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCAAGACAGTT 976
Qy      781  GTACTGTCATGTCATCTGAGACAGCTGGGATGTTGGGCAAGGTGCATGATGATGCGC 840
Db      977  GTACTGTCATGTCATCTGAGACAGCTGGGATGTTGGGCAAGGTGCATGATGATGCGC 1036
Qy      841  GGTGGAGCCTTTATATCAAGGTAATTTTCAACTTATTAAGATCTTGGGCTGGTCCA 900
Db      1037  GGTGGAGCCTTTATATCAAGGTAATTTTCAACTTATTAAGATCTTGGGCTGGTCCA 1096
Qy      901  AAGAGACTCTGCGGCTGTGCTCTGGACTGCAGAGAAAGAGTGGATGTTGCTCTTC 960
Db      1097  AAGAGACTCTGCGGCTGTGCTCTGGACTGCAGAGAAAGAGTGGATGTTGCTCTTC 1156
Qy      961  CAGTATTCAGTTTACAGAGGTAATTTTCAACTTATTAAGATCTTGGGCTGGTCCA 1020
Db      1157  CAGTATTCAGTTTACAGAGGTAATTTTCAACTTATTAAGATCTTGGGCTGGTCCA 1216
Qy      1021  GCAGGAACCTTCTTAACCACTGGGCTGCAATTTCACTGCAAGTAAAGAGGCAAGGCGCATC 1080
Db      1217  GCAGGAACCTTCTTAACCACTGGGCTGCAATTTCACTGCAAGTAAAGAGGCGCGCATC 1276
Qy      1277  ATGGAGAGGTTTATGAGCTGCTGAGCCCTCAATATCACTCAGTCTCTAGAGCATGGA 1336
Db      1141  GAAGGAGACAGATCACTTTTGGATCCAAGCTGAGTGCCTGAGCAAGTCTTACTTGAT 1200
Qy      1337  GAAGGAGACAGATCACTTTTGGATCCAAGCTGAGTGCCTGAGCAAGTCTTACTTGAT 1396
Db      1201  GACTTATTCAGTATTTCTCTTCCATCACTCCCAAGGAGACACCATATCAGTCACTGAT 1260
Qy      1397  GACTTATTCAGTATTTCTCTTCCATCACTCCCAAGGAGACACCATATCAGTCACTGAT 1456
Db      1261  CCAAGACAGATGAATGTTGCTGCTGCTTTGGGCTGTTGTTCTTATGTTGTCAGAC 1320
Qy      1457  CCAAGACAGATGAATGTTGCTGCTGCTTTGGGCTGTTGTTCTTATGTTGTCAGAC 1516

```

```

Qy      1321  ATGGAAGAAATGCTGCTAGTCC 1344
Db      1517  ATGGAAGAAATGCTGCTAGTCC 1540

RESULT 4
BC020689
LOCUS      BC020689      1928 bp      mRNA      linear      PRI 22-JAN-2002
DEFINITION Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418
IMAGE:4251802, mRNA, complete cds.
ACCESSION  BC020689
VERSION     BC020689.1  GI:18088383
KEYWORDS   MGC.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1928)
            Strausberg, R.
            Direct Submission
            Submitted (03-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
            NIH-MGC Project URL: http://mgc.nci.nih.gov
            Contact: MGC help desk
            Email: cgabs-remail.nih.gov
            Tissue Procurement: ATCC/DCT/DTP
            CDNA Library Preparation: CLONTECH Laboratories, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (HLNU)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site: http://www.shgc.stanford.edu
            Contact: (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
            R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/HLNU at: http://image.llnl.gov
Series: IRAL Plate: 35 Row: f Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7706386.

FEATURES
    source
        1..1928
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="LocusID:10404"
            /db_xref="taxon:9606"
            /clone="MGC:22418 IMAGE:4251802"
            /tissue_type="Skin, melanotic melanoma, high MDR."
            /clone_id="NIH_MGC_62"
            /lab_host="DH10B"
            /note="Vector: pDNR-LIB"
            157..1585
            /codon_start=1
            /product="plasma glutamate carboxypeptidase"
            /protein_id="AAH20689.1"
            /db_xref="GI:18088384"
            /translation="MKFLIPAPFGVHLSTLGSKAIKNGISKRPFEIKETIASCG
            DVAKIILNLAIVGKAONRSYERLALLVDITGPLSGKNEKRIQIOMYOLQDGLDI
            VHEPVRIPHEMERBSAVMLBPRIKIATLIGGSSIGTPPEGITAEVLVYTSDEIQ
            RASEARGLVIVNQPIYNSRTVQRTQGAVERAAKVGALASLRVSASTISPHTG
            IREYODGVPKIPACITVEDAEKMSRNASGIRIVIQKXGATYPTPDSFNVVAETG
            GSKYPEGVIVLVSGLSDSDVGGQAMDDGGAFLISWEALSLIKDLGLRPRKTLILVMT
            AEROGVGARVOYOLHKVNTSNGSLVMSAGTFLPGLQFTSEKRAIMEEYMSLL
            OPNTIQVLSHGSGTIDNFWIOAGVPEASLDDLKXFFPHSHSGDMTMDKQNV
            AAAYVAVSYVAVDMEMLPRS"

BASE COUNT  554 a      435 c      466 g      473 t
ORIGIN
Query Match      99.8%; Score 1340.8; DB 9; Length 1928;
Best Local Similarity 99.9%; Pred. No. 0;

```

Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AAGATGCACTCTCTAGAGAGACTTTTGAAGAATAAAGAAATAGCAGCTGGGA 60
DB 239 AAGATGCACTCTCTAGAGAGACTTTTGAAGAATAAAGAAATAGCAGCTGGGA 298
QY 61 GATGTTCTAAGCAATCATCACTAGCTGTTTATGTAAGCCAGAAAGATCTTAT 120
DB 299 GATGTTCTAAGCAATCATCACTAGCTGTTTATGTAAGCCAGAAAGATCTTAT 358
QY 121 GAGGATGTCAGCTTCTGTTGATGATCTGTGAGACCCAGCTGATGCTCCAGAACCTTA 180
DB 359 GAGGATGTCAGCTTCTGTTGATGATCTGTGAGACCCAGCTGATGCTCCAGAACCTTA 418
QY 181 GAAAAAGCCATCCAAATTTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTTAC 240
DB 419 GAAAAAGCCATCCAAATTTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTTAC 478
QY 241 CTGAGGCAAGTGAATATCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAG 300
DB 479 CTGAGGCAAGTGAATATCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAG 538
QY 301 CCAAGAAATTCATAGATAGCCATCTGGGCTTGGCAGCAGCATTTGGGACTCTCCAGAA 360
DB 539 CCAAGAAATTCATAGATAGCCATCTGGGCTTGGCAGCAGCATTTGGGACTCTCCAGAA 598
QY 361 GGCATTACAGAGAAAGTTCTGTTGATGACCTTTTCATGATGATGAGAGAGGGCTTCA 420
DB 599 GGCATTACAGAGAAAGTTCTGTTGATGACCTTTTCATGATGATGAGAGAGGGCTTCA 658
QY 421 GAAGCAAGAGAGAAAGTTGTTTATATACCAACCTTACATCACTACTCAAGAGCCGTG 480
DB 659 GAAGCAAGAGAGAAAGTTGTTTATATACCAACCTTACATCACTACTCAAGAGCCGTG 718
QY 481 CATATCCGAACGACGGGGGGGGTGGAAAGCTGCGCAAGGTGGGGCTTTGGCATCTCTCAT 540
DB 719 CATATCCGAACGACGGGGGGGGTGGAAAGCTGCGCAAGGTGGGGCTTTGGCATCTCTCAT 778
QY 541 CGATCCGTTGCTCTCTCATCTAGCTCTGACAGCTGACAGGATTTCCAGAAATACAGAT 600
DB 779 CGATCCGTTGCTCTCTCATCTAGCTCTGACAGGATTTCCAGAAATACAGAT 838
QY 601 GGGGTGCCCCAAATTTCCAGAGCTGTATTAAGGTGGAAGATGAGAAATGATGTCAAGA 660
DB 839 GGGGTGCCCCAAATTTCCAGAGCTGTATTAAGGTGGAAGATGAGAAATGATGTCAAGA 898
QY 661 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGCTTAAGATGGGGGCAAAAGACTTACCA 720
DB 899 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGCTTAAGATGGGGGCAAAAGACTTACCA 958
QY 721 GATACGATTCCTTCAACAGCTGTAGCAGAGATCTGAGGAGCAATATCCAGAAACAGATT 780
DB 959 GATACGATTCCTTCAACAGCTGTAGCAGAGATCTGAGGAGCAATATCCAGAAACAGATT 1018
QY 781 GTACTGTCAGTGAACATCTGAGCAGCTGGGATTTGGGCGAGGTGCCATGATGATGAGC 840
DB 1019 GTACTGTCAGTGAACATCTGAGCAGCTGGGATTTGGGCGAGGTGCCATGATGATGAGC 1078
QY 841 GGTGAGACCTTTATATCATGGGAAGCACTCTCATTTAAGATCTTGGGCTGCTGCA 900
DB 1079 GGTGAGACCTTTATATCATGGGAAGCACTCTCATTTAAGATCTTGGGCTGCTGCA 1138
QY 901 AAGAGAGCTGCGGGGCTGTGCTCTGACCTGCAAGAAAGAGTTGATGTTGCTGCTTC 960
DB 1139 AAGAGAGCTGCGGGGCTGTGCTCTGACCTGCAAGAAAGAGTTGATGTTGCTGCTTC 1198
QY 961 CAGTATTATGATACCAAGATTAATTTCAACTACAGTCTGTGATGATGATGATGATGATGAT 1020
DB 1199 CAGTATTATGATACCAAGATTAATTTCAACTACAGTCTGTGATGATGATGATGATGATGAT 1258
QY 1021 GCAGGAACCTTTTACCACTGGGCTGCAATTCATGCGCAGTGAAGAGCCAGGGCCATC 1080
DB 1259 GCAGGAACCTTTTACCACTGGGCTGCAATTCATGCGCAGTGAAGAGCCAGGGCCATC 1318

```

```

QY 1081 ATGAGAGAGTTATGAGACCTGTGACAGCCCTCAATATCACTGAGGCTGAGCATGGA 1140
DB 1319 ATGAGAGAGTTATGAGACCTGTGACAGCCCTCAATATCACTGAGGCTGAGCATGGA 1378
QY 1141 GAAAGGACAGACATCAACTTTTGATCCAGCTGAGAGTGGCTGGAGCCAGCTTATGAT 1200
DB 1379 GAAAGGACAGACATCAACTTTTGATCCAGCTGAGAGTGGCTGGAGCCAGCTTATGAT 1438
QY 1201 GACTTATACAAATTTTCTTCTTCCATCATCTCCAGGAGAACACATGATCTGATGAT 1260
DB 1439 GACTTATACAAATTTTCTTCTTCCATCATCTCCAGGAGAACACATGATCTGATGAT 1498
QY 1261 CCAAGCAGATGATGTTGCTGCTGTTGGGCTGTTGTTCTTATGTTTGGAGAC 1320
DB 1499 CCAAGCAGATGATGTTGCTGCTGTTGGGCTGTTGTTCTTATGTTTGGAGAC 1558
QY 1321 ATGAGAAATGCTGCTAGTCC 1344
DB 1559 ATGAGAAATGCTGCTAGTCC 1582

```

```

RESULT 5
AF107834 1472 bp mRNA linear PRI 12-JUL-1999
LOCUS AF107834
DEFINITION Homo sapiens clone LCHI aminopeptidase mRNA, complete cds.
ACCESSION AF107834
VERSION AF107834.1 GI:5442029
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
1 (bases 1 to 1472)
Liu, C.H., Lin, B.Y. and Chang, L.Y.
Cloning of the human aminopeptidase gene
Unpublished
REFERENCE
2 (bases 1 to 1472)
Liu, C.H., Lin, B.Y. and Chang, L.Y.
Direct Submission
Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
Sinica, RM 433, 128, Yen-Chiun-Yuan Road SEC 2, Taipei 11529, Taiwan
JOURNAL
FEATURES
source
1..1472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LCHI"
/tissue_type="liver"
/dev_stage="fetus"
54..1472
/feature="secretory protein"
/codon_start=1
/product="aminopeptidase"
/protein_id="A043214.1"
/db_xref="GI:5442030"
/translation="MKFLIPAFPGVHLLSCGKAI CKNGISKRTPEETKEIASCG
DVAKIINLAIVYKAQNRSTERYLALIVDTVPPLSSKINLEKAIQIYQWLQDDLEK
VHLEPVRIPEHMERGESAVLEPRIRKIALIGSSISITGPBEITAEVLVTSFDEIQ
RASEBARKIVVINOPIYINSRVQRTQGAANAAGVLAASLIRSVASFSYSPHG
IOEYOGVPRKIPACITVEDAEWMSWASHGIRIVOLKMGAKATYPTDSFNTVAEIT
GSKYPRQVVLVSHLSDMDVCGGAMDDGGAFTSWBALSLIKLGLRPRKTLVIMT
AEQGGVGAFOYIQLHKVINSNYSLYMESBAGTFLPTGLDPTFSBKAALMEVMSIL
QPLNTIQLVSHGSDTINFIQAGVPGASLLDLYRFFPHSHGDTMTVMDPRQKNV
AAVWAVVSVVAVMDMEMLPRS"

```

```

BASE COUNT 411 a 325 c 381 g 355 t
ORIGIN

```

```

Query Match 99.5%; Score 1337.6; DB 9; Length 1472;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1340; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAGATGCACTCTCTAAGAGAGACTTTTGAAGAATAAAGAAATAGCAGCTGGGA 60

```

```

126 AAGATGGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATGCGAGCTGTGA 185
126 AAGATGGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATGCGAGCTGTGA 185
Qy 61 GATGTTGCTAAAGCATATCACTAGCTGTTTATGTTAAAGCCGAGACATCTTAT 120
Qy 186 GATGTTGCTAAAGCATATCACTAGCTGTTTATGTTAAAGCCGAGACATCTTAT 245
Qy 121 GAGGATTTGGCATCTCTGTTGATGATCTGTGGAACCACTGAGTGGCTCCAGAACTA 180
Qy 246 GAGGATTTGGCATCTCTGTTGATGATCTGTGGAACCACTGAGTGGCTCCAGAACTA 305
Qy 181 GAAAAAGCCATCAAAATTTATGATCAAAAACCTGAGCAAGATGGCTGAGAAAGTTTAC 240
Qy 306 GAAAAAGCCATCAAAATTTATGATCAAAAACCTGAGCAAGATGGCTGAGAAAGTTTAC 365
Qy 241 CTGAGAGCATGAGAAATTCCTCCACTGAGAGAGGGAGAAAGATCACTGTGTGAG 300
Qy 366 CTGAGAGCATGAGAAATTCCTCCACTGAGAGAGGGAGAAAGATCACTGTGTGAG 425
Qy 301 CCAAGAAATTCATTAAGATAGCATCTGGGCTGTTGGACAGCATTTGGGCTCTCCAGAA 360
Qy 426 CCAAGAAATTCATTAAGATAGCATCTGGGCTGTTGGACAGCATTTGGGCTCTCCAGAA 485
Qy 361 GGCATTAACAGCAGAAATTTCTGTGTGATCACTTTGATGAACTGAGAGAGGCTTCA 420
Qy 486 GGCATTAACAGCAGAAATTTCTGTGTGATCACTTTGATGAACTGAGAGAGGCTTCA 545
Qy 421 GAACCAAGAGGAGAAATTTGTTGTTTATACCAACTTATCACTACTCAAGACGGTG 480
Qy 546 GAACCAAGAGGAGAAATTTGTTGTTTATACCAACTTATCACTACTCAAGACGGTG 605
Qy 481 CAATACGGAAGCGAGGGGGCGGTGGAAGCTCAAGGAGGGGGCTTTGGCATCTCTCAT 540
Qy 606 CAATACGGAAGCGAGGGGGCGGTGGAAGCTCAAGGAGGGGGCTTTGGCATCTCTCAT 665
Qy 541 CGATCGGTGGCTCTCTCTCATCTACAGTCTTCAACAGGATTTTCAAGATACAGAT 600
Qy 666 CGATCGGTGGCTCTCTCTCATCTACAGTCTTCAACAGGATTTTCAAGATACAGAT 725
Qy 601 GGCCTGCCCAAAATTTCCAAAGCTGTATTCGGTGAAGATGCAAGATATGTCAAGA 660
Qy 726 GGCCTGCCCAAAATTTCCAAAGCTGTATTCGGTGAAGATGCAAGATATGTCAAGA 785
Qy 661 ATGGCTCTCATAGGATCAAAATTTGTCATTCAGCTAAAGATGGGGCAAAAGCTTACCA 720
Qy 786 ATGGCTCTCATAGGATCAAAATTTGTCATTCAGCTAAAGATGGGGCAAAAGCTTACCA 845
Qy 721 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGACAAATATCCAGAAAGAT 780
Qy 846 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGACAAATATCCAGAAAGAT 905
Qy 781 GATACGATTCCTTCAACACTGTAGCAGAGATTTGGGCGGGTCCATGATGATGGC 840
Qy 906 GATACGATTCCTTCAACACTGTAGCAGAGATTTGGGCGGGTCCATGATGATGGC 965
Qy 841 GGTGAGACCTTTATATATATGAGGAGCACTCTCATTTTAAAGATCTTGGCTGCTCCA 900
Qy 966 GGTGAGACCTTTATATATATGAGGAGCACTCTCATTTTAAAGATCTTGGCTGCTCCA 1025
Qy 901 AAGAGACTCTGCGGCTGTGCTCTGAGCTGAGAGAGAAAGATGATGATGCTTCC 960
Qy 1026 AAGAGACTCTGCGGCTGTGCTCTGAGCTGAGAGAGAAAGATGATGATGCTTCC 1085
Qy 961 CAGATATATCAGTTACACAGAGTAAATTTTCCAACTACAGTCTGTGATGAGATCTGAC 1020
Qy 1086 CAGATATATCAGTTACACAGAGTAAATTTTCCAACTACAGTCTGTGATGAGATCTGAC 1145
Qy 1021 GCAAGAACTTTCTTAACCACTGGGCTGCAATTCATGCGAGTGAAGAGGCGAGGCTATC 1080
Qy 1146 GCAAGAACTTTCTTAACCACTGGGCTGCAATTCATGCGAGTGAAGAGGCGAGGCTATC 1205
Qy 1081 ATGAGAGAGTTATGAGCTGTGCTGAGCCCTCAATATCACTCAGTCTGAGCATGGA 1140

```

```

Db 1206 ATGAGAGAGGTTATGAGCTGCTGAGAGCCCTCAATATCACTCAGTCTGAGCATGGA 1265
Qy 1141 GAAGGAGCAGACATCAACTTTTGGATCCAAAGCTGAGAGTGGCTGAGCAGCTTACTTAT 1200
Db 1266 GAAGGAGCAGACATCAACTTTTGGATCCAAAGCTGAGAGTGGCTGAGCAGCTTACTTAT 1325
Qy 1201 GACTTATACAGATATTTCTTCCATCACTCCCAAGAGAGACACATGATGATGATGAT 1260
Qy 1326 GACTTATACAGATATTTCTTCCATCACTCCCAAGAGAGACACATGATGATGATGAT 1385
Qy 1261 CCAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1386 CCAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1445
Qy 1321 ATGAGAAATTCCTGCTTACTGCTCC 1344
Db 1446 ATGAGAAATTCCTGCTTACTGCTCC 1469

```

RESULT 6
BD127520 1860 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION
Primer for synthesizing full-length cDNA and use thereof.
ACCESSION
BD127520
VERSION
BD127520.1 GI:23222465
KEYWORDS
JP 2002017375-A/2951.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
Oka, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Makamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
1 (bases 1 to 1860)
TITLE
Primer for synthesizing full-length cDNA and use thereof
JOURNAL
Patent: JP 2002017375-A 2951 22-JAN-2002;
HELIX RESEARCH INSTITUTE
COMMENT
OS Homo sapiens (human)
PN JP 2002017375-A/2951
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OKA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YUJI KAWAI, AI MAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUO OTSUKI, HISASHI KOGA

FEATURES
source
1. 1860
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 515 a 423 c 451 g 471 t
ORIGIN
Query Match 99.4%; Score 1336; DB 6; Length 1860;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1339; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATGCGAGCTGTGA 60
Db 201 AAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATGCGAGCTGTGA 260
Qy 61 GATGTTGCTAAAGCATATCACTAGCTGTTTATGTTAAAGCCGAGACATCTTAT 120
Db 261 GATGTTGCTAAAGCATATCACTAGCTGTTTATGTTAAAGCCGAGACATCTTAT 320

QY 121 GAGCGATTGGCACTTGTGTTGATCTGTGAGACCCAGACTGAGTGGCTCCAGAACCTTA 180
 Db 321 GAGCGATTGGCACTTGTGTTGATCTGTGAGACCCAGACTGAGTGGCTCCAGAACCTTA 380
 QY 181 GAAAAACCCATCCAAATTAATGACCAAAACCTGACAGCAAGTGGGCTGGAGAAAGTTAC 240
 Db 381 GAAAAACCCATCCAAATTAATGACCAAAACCTGACAGCAAGTGGGCTGGAGAAAGTTAC 440
 QY 241 CTGGAGCCAGTGAAGTAATCCCACTGGAGAGAGGAGAGAAATCAGCTGTATGCTGGAG 300
 Db 441 CTGGAGCCAGTGAAGTAATCCCACTGGAGAGAGGAGAGAAATCAGCTGTATGCTGGAG 500
 QY 301 CCAAGAAATTCATAAGATAGCCATCTGGGCTTGGACAGACATTTGGGACTCTTCCAGAA 360
 Db 501 CCAAGAAATTCATAAGATAGCCATCTGGGCTTGGACAGACATTTGGGACTCTTCCAGAA 560
 QY 361 GGCATTACAGCAAGATTTCTGTGTGACCTTTTCATGTAAGTGCAGAGAAAGGCTTCA 420
 Db 561 GGCATTACAGCAAGATTTCTGTGTGACCTTTTCATGTAAGTGCAGAGAAAGGCTTCA 620
 QY 421 GAAGCAAGAGGAGGATTTGTTTATTAACCACTTACATCACTACTCAAGAGAGGCTG 480
 Db 621 GAAGCAAGAGGAGGATTTGTTTATTAACCACTTACATCACTACTCAAGAGAGGCTG 680
 QY 481 CAATACCGAAGCAGAGGGGCGGTGGAAGCTCCAAAGTGGGGCTTTGGCATCTTCAATT 540
 Db 681 CAATACCGAAGCAGAGGGGCGGTGGAAGCTCCAAAGTGGGGCTTTGGCATCTTCAATT 740
 QY 541 CGATCCGTGGCTCTCTTCTCATCTACAGTCTTCAACAGGTTATTCAGAAATACAGAGAT 600
 Db 741 CGATCCGTGGCTCTCTTCTCATCTACAGTCTTCAACAGGTTATTCAGAAATACAGAGAT 800
 QY 601 GGGCGGCCCAAAATTCACAGCGCTGATTAAGGAGAGAGTGAAGTGAAGTATGTCAGA 660
 Db 801 GGGCGGCCCAAAATTCACAGCGCTGATTAAGGAGAGAGTGAAGTGAAGTATGTCAGA 860
 QY 861 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGTCAAGTGGGGCGCAAGACCTACCA 720
 Db 861 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGTCAAGTGGGGCGCAAGACCTACCA 920
 QY 721 GATATCTGATTTCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTT 780
 Db 921 GATATCTGATTTCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTT 980
 QY 781 GATATCTGATTTCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTT 840
 Db 981 GATATCTGATTTCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTT 1040
 QY 841 GGTGGAGCTTTATATCATGAGGAAGCACTTCACTTATTAAGATCTTGGGCTGGCTCA 900
 Db 1041 GGTGGAGCTTTATATCATGAGGAAGCACTTCACTTATTAAGATCTTGGGCTGGCTCA 1100
 QY 901 AAGAGACTCTGCGGCTGTGCTCTGACTGACAGAGAAACAAGGTGAGTTGTCCTTC 960
 Db 1101 AAGAGACTCTGCGGCTGTGCTCTGACTGACAGAGAAACAAGGTGAGTTGTCCTTC 1160
 QY 961 CAGTATTAATCACTTACAGAAAGTAATTTTCCAACTACAGTGTGATGAGAGTCTGAC 1020
 Db 1161 CAGTATTAATCACTTACAGAAAGTAATTTTCCAACTACAGTGTGATGAGAGTCTGAC 1220
 QY 1021 GCAGGAACCTTCTTATACCACTGGGCTGCAATTCATCTGGCACTGAGAAAGCCAGGCCATC 1080
 Db 1221 GCAGGAACCTTCTTATACCACTGGGCTGCAATTCATCTGGCACTGAGAAAGCCAGGCCATC 1280
 QY 1081 ATGAGAGAGTTATGAGCTGCTGACGCCCTCAATATCACTCAGGTCTGAGCCATGGA 1140
 Db 1281 GTGAGAGAGTTATGAGCTGCTGACGCCCTCAATATCACTCAGGTCTGAGCCATGGA 1340
 QY 1141 GAAGGAGAGATCACTTTTGGATCCAGCTGAGAGTGGCTGGAGCCAGTCTACTTAT 1200
 Db 1341 GAAGGAGAGATCACTTTTGGATCCAGCTGAGAGTGGCTGGAGCCAGTCTACTTAT 1400
 QY 1201 GACTTAATCAAGTATTTCTTCTCATCACTCCACGAGAGACCATGATGTATGAT 1260

Db 1401 GACTTAATCAAGTATTTCTTCTCATCACTCCACGAGAGACCATGATGTATGAT 1460
 QY 1261 CCAAGAGATGATATTTGCTGCTGCTGCTTGGGCTTGTATGTTTGTGACAC 1320
 Db 1461 CCAAGAGATGATATTTGCTGCTGCTGCTTGGGCTTGTATGTTTGTGACAC 1520
 QY 1321 ATGAGAGAAATGCTGCTAGGCTC 1344
 Db 1521 ATGAGAGAAATGCTGCTAGGCTC 1544

RESULT 7
 AK075132
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ90651 f18, clone PLACE1004482, moderately similar to Rattus norvegicus hematopoietic lineage switch 2 related protein (Hlzf-rp) mRNA.
 AK075132
 ACCESSION AK075132.1 GI:22761022
 VERSION o1igo capping; f18 (full insert sequence).
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 1 Isegai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T., Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S., Kawai, Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kojima, S., Nagahara, K., Masuno, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Nimomiva, K.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 1860)
 Isegai, T. and Otsuki, T.
 Direct Submission
 Submitted (25-MAR-2002) Takao Isegai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genome@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).
 Location/Qualifiers
 1..1860

COMMENT
 FEATURES
 source
 1..1860
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1004482"
 /tissue_type="Placenta"
 /clone_id="PLACE1"
 /note="Cloning vector: pME18SFL3"
 129..1547
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC11423.1"
 /db_xref="GI:22761023"
 /translation="MKRLIFAPRGVHLLSLCSKAI CRNGISKRTFEIEEIASCG DVAKATINLVYGAQRSTYRLALVDIVGPRUSGSKNLEKAIQINQDQGLEK VHLPEPVI PHEMERGESAVLLEPRILHIALGLSSIGTPEGLTAEVLVTSDELQ RRASEARGLVNNQPIYNSRTVOYRTOGAVEAKGALASIRSVASFISYPHTG IOEXODVPKIPACITVEDAEEMMSRMAHGIRKIVIOKGAATYPTDSFNVAEIT GSKYRQEVAVVSGHIDSMVDYGQAMDDGGAFISWELSLIKDLGLPKKTLRLVMT AERQYGAFOYVOYHAKVINSYSLVWESAGTFLPGLGLOFTGSEKRAIYAEVMSL QPLATIVLHSGSTIDINFWIQGVNPASLDDLVKYFFPHSHGDTMTWDPKQNV AAAYWAVSVVAVDMEEMLPSS"

BASE COUNT 515 a 423 c 451 g 471 t
 ORIGIN

Query Match 99.4%; Score 1336; DB 9; Length 1860;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1339; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 AAGATGCGATCTCTTAAGAGAGCTTTTGAAGAAATTAAGAAATGAGGCTGTGA 60
 201 AAGATGCGATCTCTTAAGAGAGCTTTTGAAGAAATTAAGAAATGAGGCTGTGA 260
 61 GATGTGCTTAAGAGAGATCAACCTAGCTGTTTATGTAAGAGGAGAGAGATCTAT 120
 261 GATGTGCTTAAGAGAGATCAACCTAGCTGTTTATGTAAGAGGAGAGAGATCTAT 320
 121 GAGCGATTGGCACTTCTGGTTGATCTGTTGAGAGGAGAGAGAGAGAGAGAGAG 180
 321 GAGCGATTGGCACTTCTGGTTGATCTGTTGAGAGGAGAGAGAGAGAGAGAGAG 380
 181 GAAAAAGCCATCCAAATTAATGTAACAAAAGCTGACAGCAAGATGGGCTGAGAAAGTTAC 240
 381 GAAAAAGCCATCCAAATTAATGTAACAAAAGCTGACAGCAAGATGGGCTGAGAAAGTTAC 440
 241 CTGAGAGCGATGAGAAATCCCACTGGAGAGAGGAGAGAGAGATGAGCTGATGAG 300
 441 CTGAGAGCGATGAGAAATCCCACTGGAGAGAGGAGAGAGAGATGAGCTGATGAG 500
 301 CCAGAAATTCATAGATAGCCATCTGGGCTTGGAGAGAGAGAGAGAGAGAGAGAG 360
 501 CCAGAAATTCATAGATAGCCATCTGGGCTTGGAGAGAGAGAGAGAGAGAGAGAG 560
 361 GGCATTACAGCAAGATCTGGTGTGACCTCTTTGATGAATCTGAGAGAGAGAGAG 420
 561 GGCATTACAGCAAGATCTGGTGTGACCTCTTTGATGAATCTGAGAGAGAGAGAG 620
 421 GAAGCAAGAGAGAGATGTTGTTTATTAACCAACCTTACATACCTACAGAGAGAG 480
 621 GAAGCAAGAGAGAGATGTTGTTTATTAACCAACCTTACATACCTACAGAGAGAG 680
 481 CAATAACGAG 540
 681 CAATAACGAG 740
 541 CGATCCGAG 600
 741 CGATCCGAG 800
 601 GGCAG 660
 801 GGCAG 860
 661 ATGCTCTCTCATGAGATCAAAATTTGATTCAGTAAAGATGGGAGAGAGAGAG 720
 861 ATGCTCTCTCATGAGATCAAAATTTGATTCAGTAAAGATGGGAGAGAGAGAG 920
 721 GATACCTGATCTCTTCAACACTGTAGCAGAGATCACTGGAGAGAGAGAGAGAG 780
 921 GATACCTGATCTCTTCAACACTGTAGCAGAGATCACTGGAGAGAGAGAGAGAG 980
 781 GATACCTGATCTCTTCAACACTGTAGCAGAGATCACTGGAGAGAGAGAGAGAG 840
 981 GATACCTGATCTCTTCAACACTGTAGCAGAGATCACTGGAGAGAGAGAGAGAG 1040
 841 GGTGAG 900
 1041 GGTGAG 1100
 901 AAG 960
 1101 AAG 1160
 961 CAGTATTAACAGTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 1161 CAGTATTAACAGTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1220
 1021 GCAG 1080

Db 1221 GCAG 1280
 Qy 1081 ATGAG 1140
 Db 1281 GTGAG 1340
 Qy 1141 GAAG 1200
 Db 1341 GAAG 1400
 Qy 1201 GACTTATACAGATATTTCTTTCATCACTCCAGAGAGAGAGAGAGAGAGAGAG 1260
 Db 1401 GACTTATACAGATATTTCTTTCATCACTCCAGAGAGAGAGAGAGAGAGAGAG 1460
 Qy 1261 CCAG 1320
 Db 1461 CCAG 1520
 Qy 1321 ATGAG 1380
 Db 1521 ATGAG 1580

RESULT 8
 AF119386 1794 bp mRNA linear PRI 20-MAY-1999
 LOCUS AF119386
 DEFINITION Homo sapiens blood plasma glutamate carboxypeptidase precursor
 ACCESSION AF119386
 VERSION AF119386.1 GI:4877697
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1794)
 Gengras, R., Richard, C., El-Alfy, M., Morales, C.R., Potier, M. and
 Pehezhetsky, A.V.
 Purification, cDNA cloning, and expression of a new human blood
 plasma glutamate carboxypeptidase homologous to
 N-acetyl-aspartyl-alpha-glutamate
 carboxypeptidase/prostate-specific membrane antigen
 J. Biol. Chem. 274 (17), 11742-11750 (1999)
 99223495
 JOURNAL MEDLINE
 PUBMED 10206390
 REFERENCE
 AUTHORS Gengras, R., Richard, C., El-Alfy, M., Morales, C.R., Potier, M. and
 Pehezhetsky, A.V.
 TITLE Submitted (12-JUN-1999) Medical Genetics, Sainte-Justine Hospital,
 JOURNAL Montreal University, 3175 Cote Sainte-Catherine, Montreal, QU H3T
 1CS, Canada
 FEATURES
 source
 1. 1794
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8q22.2; between D8S257 and D8S270"
 1. 1794
 /gene="PGCP"
 /gene="PGCP"
 25. 1650
 /gene="PGCP"
 /EC_number="3.4.17.21"
 /function="endopeptidase"
 /note="member of M28 peptidase family; similar to
 glutamate carboxypeptidase II and to
 N-acetyl-aspartyl-alpha-glutamate carboxypeptidase
 (NAGLase); similar to sequences with Genbank Accession
 Numbers R18560 and G07404; prostate-specific membrane
 antigen (PSMA)"
 /codon_start=1


```

/product="blood plasma glutamate carboxypeptidase
precursor"
/protein_id="AA011418.1"
/db_xref="GI:4877698"
/translation="MKFLIRFPGVHVLISCSKAIKRGKISKRTFEIKELIASG
DVAKATINLAVYKAQNRYSERLALVDIVGPRISGSKNLEKAIQIMYQMLQDDLEK
VHLEPVI PHMERGEESAVMLEPRIHAIILGLSISIGTPEGITIAVLIVTDFEIQ
RRASEARGLIVVYQPYINYSRTVOYRTOGAVEAKYALASLIRSVAFSIYSPHG
IOEKDQVPIKTPACITVEDAEKMSRMAHGKIKIVLQKGAATKPTSDSFVASEIT
GSKYRQVAVLSGHLDSMDYGGAMDGGCAFTSMELSLIKDLGAPKRTLRVMT
AEEQGVGAFQYVQYHLYKVNISNYSLVNESDAGTFLPTGLDPTSEKRAIMEEMSLI
QPLNTIYVLSHGETDINFWIQAGVPASLIDLDYKFFPHSHGDTMTWDSRWML
LPLGLFLMLQITWKKCCIGPRNSKKEITSCFWPGLIGLQMTPLHITISNSSSK
HNSISCFLLSFLISKFSRKRNSPLPPT"
25..96
sig_peptide
/gene="PGCP"
157..1647
mat_peptide
/gene="PGCP"
/product="blood plasma glutamate carboxypeptidase"
205..207
/misc_feature
/note="putative; glycosylation site"
559..561
/misc_feature
/gene="PGCP"
/note="putative; glycosylation site"
1081..1083
/misc_feature
/gene="PGCP"
/note="putative; glycosylation site"
1090..1092
/misc_feature
/gene="PGCP"
/note="putative; glycosylation site"
1210..1212
/misc_feature
/gene="PGCP"
/note="putative; glycosylation site"
BASE COUNT 522 a 391 c 422 g 457 t 2 others
ORIGIN
Query Match 98.8%; Score 1327.2; DB 9; Length 1794;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1340; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 AAGATGGCATCTCTAAGAGACTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 60
Db 97 AAGATGGCATCTCTAAGAGACTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 156
QY 61 GATTTGCTAAAGCAATCATCACTTACTGTTTATGTTAAAGCCAGAACAGATCTTAT 120
Db 157 GATTTGCTAAAGCAATCATCACTTACTGTTTATGTTAAAGCCAGAACAGATCTTAT 216
QY 121 GAGGATTTGGCATCTTCTGTTGATATCTGTGGACCCAGACTGATGGCTCCAGAACCTTA 180
Db 217 GAGGATTTGGCATCTTCTGTTGATATCTGTGGACCCAGACTGATGGCTCCAGAACCTTA 276
QY 181 GAAAAAGCCATCCAAATATATGATCCAAAACTGCGAGCAAGATGGGCTGAGAAAGTTTAC 240
Db 277 GAAAAAGCCATCCAAATATATGATCCAAAACTGCGAGCAAGATGGGCTGAGAAAGTTTAC 336
QY 241 CTGAGCCAGTGAAGATATCCCACTGGAGAGGGGAGAAAGATAGCTGTGATCTGGAG 300
Db 337 CTGAGCCAGTGAAGATATCCCACTGGAGAGGGGAGAAAGATAGCTGTGATCTGGAG 396
QY 301 CCAAGATTCATTAAGATAGCCATCTCTGGGCTTTGGCAGACAGATTGGAGCTCTCCAGAA 360
Db 397 CCAAGATTCATTAAGATAGCCATCTCTGGGCTTTGGCAGACAGATTGGAGCTCTCCAGAA 456
QY 361 GGCATTACAGAGAGATTCTGTGGTGAATCTTTTCAGATGAAGTGCAGAGAAAGGAGCTCA 420
Db 457 GGCATTACAGAGAGATTCTGTGGTGAATCTTTTCAGATGAAGTGCAGAGAAAGGAGCTCA 516
QY 421 GAAGCAAGAGGAGAGATTGTTGTTTAAACCAACTTACATCACTACTCAAGAGCGGTG 480
Db 517 GAAGCAAGAGGAGAGATTGTTGTTTAAACCAACTTACATCACTACTCAAGAGCGGTG 576

```

```

QY 481 CAATACCGAAGCGAGGGGGCGGTGAAGCTGCCAAGGTGGGGCTTTGGCATCTCAT 540
Db 577 CAATACCGAAGCGAGGGGGCGGTGAAGCTGCCAAGGTGGGGCTTTGGCATCTCAT 636
QY 541 CGATTCGGTGGCTCTCTTCTTCATCTACAGTCTCTCAACAGATTTTCAAGAAATCCAGAT 600
Db 637 CGATTCGGTGGCTCTCTTCTTCATCTACAGTCTCTCAACAGATTTTCAAGAAATCCAGAT 696
QY 601 GGGGTGCCCAAAATTCACACAGCTGATTTAGGGTGAAGATCAGAAATGATGTCAGA 660
Db 697 GGGGTGCCCAAAATTCACACAGCTGATTTAGGGTGAAGATCAGAAATGATGTCAGA 756
QY 661 ATGGCTTTCATGGATCAAAATTTGTCATTCAGCTAAAGATGGGGCAAAAGACTTCCA 720
Db 757 ATGGCTTTCATGGATCAAAATTTGTCATTCAGCTAAAGATGGGGCAAAAGACTTCCA 816
QY 721 GATATGATTCCTTCAACACTGTAGAGAGATCACTGGAGCAATATTCAGAAACAGTT 780
Db 817 GATATGATTCCTTCAACACTGTAGAGAGATCACTGGAGCAATATTCAGAAACAGTT 876
QY 781 GTACTGGTCACTGAGCATCTGAGCACTGGGATGTTGGGAGGGTCCATGATGATGGC 840
Db 877 GTACTGGTCACTGAGCATCTGAGCACTGGGATGTTGGGAGGGTCCATGATGATGGC 936
QY 841 GGTGAGGCTTTTATCATAGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCA 900
Db 937 GGTGAGGCTTTTATCATAGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCA 996
QY 901 AAGAGACTCTGGGGCTGGTCTGTGAGCTGCAAGAAAGACAGGTGAGTTGGTCCCTTC 960
Db 997 AAGAGACTCTGGGGCTGGTCTGTGAGCTGCAAGAAAGACAGGTGAGTTGGTCCCTTC 1056
QY 961 CAGTATTAATAGTTACACAAAGTAAATATTTCCAACTACAGTCTGGATGAGTCTGAC 1020
Db 1057 CAGTATTAATAGTTACACAAAGTAAATATTTCCAACTACAGTCTGGATGAGTCTGAC 1116
QY 1021 GCAGGAACCTTTTACCACTGGGCTGCAATTCATCTGACAGTAAAGAGCCAGAGCCATC 1080
Db 1117 GCAGGAACCTTTTACCACTGGGCTGCAATTCATCTGACAGTAAAGAGCCAGAGCCATC 1176
QY 1081 ATGAGAGAGTTATGAGCTGTGAGAGCCCTCAATATCACTGAGTCTGAGCCATGGA 1140
Db 1177 ATGAGAGAGTTATGAGCTGTGAGAGCCCTCAATATCACTGAGTCTGAGCCATGGA 1236
QY 1141 GAAGGACACACATCAACTTTTGGATTCAGAGCTGAGAGTCTGAGCACTTACTTAT 1200
Db 1237 GAAGGACACACATCAACTTTTGGATTCAGAGCTGAGAGTCTGAGCACTTACTTAT 1296
QY 1201 GACTTATACAGATTTCTTCTTCATCACTCCAGGAGACACATGACTGATGAT 1260
Db 1297 GACTTATACAGATTTCTTCTTCATCACTCCAGGAGACACATGACTGATGAT 1356
QY 1261 CCAAGCAGATGATGTTGCTGCTGTGTTGGGCTGTTTCTTATGTTGTGAGAC 1320
Db 1357 CC-CAGCAGATGATGTTGCTGCTGTGTTGGGCTGTTTCTTATGTTGTGAGAC 1415
QY 1321 ATGGAAGAAATGCTGCTAGGTC 1344
Db 1416 ATGGAAGAAATGCTGCTAGGTC 1439

```

```

RESULT 9
BC037067 1796 bp mRNA linear ROD 16-APR-2003
LOCUS
DEFINITION
MUS musculus plasma glutamate carboxypeptidase, mRNA (cDNA clone
MGC:46857 IMAGE:4989536), complete cds.
ACCESSION
BC037067
VERSION
BC037067.1 GI:22477497
KEYWORDS
MGC.
SOURCE
MUS musculus (house mouse)
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

REFERENCE
AUTHORS
1 (baaes 1 to 1796)
Strausberg,R.L., Fellnagold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Wax,S.I., Wang,J., Hsieh,F.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Cavaant,T.L.,
Scheer,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S., Locuelli,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hall,K.S.W.,
Vollard,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahney,J., Heltom,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Kitzynski,M.I., Skalka,U., Smalins,D.E.,
Schnerich,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
MEDLINE
22388257
12477932
2 (baaes 1 to 1796)
Strausberg,R.
Direct Submission
Submitted (23-Aug-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgi.mgi.nih.gov>
Contact: MGC help desk
Email: cgabbe-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breem,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,M., Ho,S.-L., Karling,E., Kwong,P., Latic,P., Legaspi,R.,
Madduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantipop,S., Thomas,P.D., Touchman,J.W.,
Tsurguen,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

REMARK
COMMENT
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAX Plate: 81 Row: 9 Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 9055233.
Location/Qualifiers
1. 1796
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:46857 IMAGE:4989536"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_lib="NCI CGAP_Co24"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1796
/gene="pgcp-pending"
/note="synonyms: HUS2, 1190003P12R1k"
/db_xref="locusID:54381"
/db_xref="MGI:1889205"
63. 1475

FEATURES
SOURCE
gene
CDS

/codon_start=1
/product="Plasma glutamate carboxypeptidase"
/protein_id="AAH37067.1"
/db_xref="GI:22477498"
/db_xref="LocusID:54381"
/translacion="MRSFLPLFVTHLALSGKAVFNKVSQRTFREIKEIANEDV
AKALINAVGKYONRSYERILGLIVDIPRLSSKNLEKAIQIMYONLODDLENYH
LEOYRIPIHMERSESAVMEPRILKMAIILIGSSIGTPPGITAEVAVSFDLQR
ASERAGKITIYNOPTGYEKTQYRVGAGVAAKVAASVLSIOSVASFSTISPTGIO
KYODGVKITIYNOPTGYEKTQYRVGAGVAAKVAASVLSIOSVASFSTISPTGIO
MYPEEVLVSGHLSMDVGGALDDGGATISWALSVLQGLRPRTLRVLATRE
EQGIGISQYVELHKANISKYSLSVWEDSGTLPETGLQPTGSDVAAKIMEVNLDP
LNTVKTYSNGBDINFIQAGVGSALRDLVLYKFFPHSHGDTMTMPDKONVAA
AVMAVVAVDMDMLPRS"

BASE COUNT 526 a 372 c 418 g 480 t
ORIGIN
Query Match 80.8%; Score 1086.4; DB 10; Length 1796;
Best Local Similarity 88.0%; Pred. No. 8.3e-298;
Matches 1183; Conservative 0; Mismatches 161; Indels 0; Gaps 0;
1 AAGATGCGATCTCTTAAGAGACTTTGAAGAAATTAAGAAGAAATAGCCAGCTGCGA 60
129 AAGATGCGTGTCTTCAGCGAACATTTGAGAAATTAAGAAGAAATAGCCAGCTATGA 188
61 GATGTTGCTAAGCAATCATCAACCTAGCTGTTATAGTAAAGCCAGAACAGATCTAT 120
189 GATGTTGCTAAGCAATCATCAACCTAGCTGTTATAGTAAATACCAAGACCGGCTTAT 248
121 GAGCGATTGCGACTTCTGTGTGATATCTGAGACCCAGACTAGTGGCTCCAGAACCTTA 180
249 GAGGTTTGGGACTTCTAGTTGATATCTGAGACCCAGACTAGTGGCTCTTAAGAACTTA 308
181 GAAAGAGCATCCAAATTTATGATACCAAACTCTGACGAAGATGGCTGAGAAAGTTTAC 240
309 GAGAAAGTATTCAAATCATGATACCAAACTCTGACGAAGATGGCTGAGAAAGTTTAC 368
241 CTGAGGCGAGGAGAAATACCCACTGAGGAGAGGAGAGAGATGAGTGGTGGTCTGAG 300
369 CTGAGGAGAGGAGAAATACCCACTGAGGAGAGGAGAGAGATGAGTGGTGGTCTGAG 428
301 CCAAGAAATTCATTAAGATAGCATCTGAGTCTTGGACAGCATTTGGAGACTCTCCAGAA 360
429 CCGGAATTCATTAAGATAGCATCTGAGTCTTGGACAGCATTTGGAGACTCTCCAGAA 488
361 GGCATTACAGAGAAAGTTGTGTGTGATGATCTTTCATGATGATGATGATGATGATGAT 420
489 GGCATTACAGAGAAAGTTGTGTGTGATGATCTTTCATGATGATGATGATGATGATGAT 548
421 GAAGCAAGAGGAGATGTTGTTTATTAACCAACTTATCACTCACTCACTCACTCACTCA 480
549 GAAGCAAGAGGAGATGTTGTTTATTAACCAACTTATCACTCACTCACTCACTCACTCA 608
481 CAATACCGAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
609 CAGTACCGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 668
541 CGATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
669 CAATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 728
601 GGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
729 GGTGTGCGCAAGATTCACAGCGCTGTATACAGTAAAGATGAGAAATGATATCTTCA 788
661 ATGCGCTTCTCATGGAGATCAAAATTTGATTCAGCTAAAGATGAGGAGGAGGAGGAGGAG 720
789 ATGCGCTTCTCATGGAGATCAAAATTTGATTCAGCTAAAGATGAGGAGGAGGAGGAGGAG 848
721 GATATGATTTCTTCAACACTGTAGCAGAGATCACTGGAGGAGGAGGAGGAGGAGGAGGAG 780
849 GATATGATTTCTTCAACACTGTGTGAGAGATCACTGGAGGAGGAGGAGGAGGAGGAGGAG 908

```

QY 781 GTACTGTCAGTGAACATCTGACACGCTGGATGTTGGGAGGGTGCATGATGATGC 840
DB 909 GTCTGCTGAGTGGACATTTGGACAGCTGGGATGTTGGGAGGGTGCATGATGATGC 968
QY 841 GGTGGAGCCTTTATATATGAGGAAAGCACTCTCACTTATTAAGATCTTGGGCTGCCTCA 900
DB 969 GGTGGAGCCTTTATATATGAGGAAAGCACTCTCACTTATTAAGATCTTGGGCTGCCTCA 1028
QY 901 AAGAAGACTGCGGGCTGGTGTCTGACATGCGAAGAAAGAGAGAGATGAGTGGCTTC 960
DB 1029 AAGAAGACTGCGGGCTGGTGTCTGACATGCGAAGAAAGAGAGAGATGAGTGGCTTC 1088
QY 961 CAGTATATCAGTTACACAAGATTAATATTTCCAACTACAGTCTGTGATGAGTCTGAC 1020
DB 1089 CAGTATATCAGTTACACAAGATTAATATTTCCAACTACAGTCTGTGATGAGTCTGAC 1148
QY 1021 GCAGGAACCTTTCTTACCCACTGGGCTGCAATTCACCTGACAGTGAAGAGCCAGGCTATC 1080
DB 1149 TCAGGAACCTTTCTTACCCACTGGGCTGCAATTCACCTGACAGTGAAGAGCCAGGCTATC 1208
QY 1081 ATGAGGAGGTTATGAGCCTGCTGCGACCCCTCAATATCACTGAGTCTGAGCCATGGA 1140
DB 1209 ATGAGGAGGTTATGAGCCTGCTGCGACCCCTCAATATCACTGAGTCTGAGCCATGGA 1268
QY 1141 GAAGGACACAGACATCACTTTTGGATCCAAAGCTGAGTGCCTGGAGCCAGTCTACTGAT 1200
DB 1269 GAAGGACACAGACATCACTTTTGGATCCAAAGCTGAGTGCCTGGAGCCAGTCTACTGAT 1338
QY 1201 GACTTATACAGTATTTCTTCTTCATCACTCCACGAGAGACACCATGACTGTGATGAT 1260
DB 1329 GACTTATACAGTATTTCTTCTTCATCACTCCACGAGAGACACCATGACTGTGATGAT 1388
QY 1261 CCAAGAGATGATGATGTTGCTGCTGCTGTTGGGCTGTGTTTATATGTTGTCAGAC 1320
DB 1389 CCAAGAGATGATGATGTTGCTGCTGCTGTTGGGCTGTGTTTATATGTTGTCAGAC 1448
QY 1321 ATGAGAAGATGCTGCTAGGTCC 1344
DB 1449 ATGAGAAGATGCTGCTAGGTCC 1472

```

RESULT 10
AF107835 1596 bp mRNA linear ROD 12-JUN-1999
LOCUS AF107835
DEFINITION Mus musculus c1one LCH-M1 aminopeptidase mRNA, complete cds.
ACCESSION AF107835
VERSION AF107835.1 GI:5442031
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.
TITLE Cloning of the mouse aminopeptidase gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1596)
AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
Sinica, RM 433, 128, Yen-Chiun-Yuan Road SEC 2, Taipei 11529, Taiwan

FEATURES
SOURCE
1..1596
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="LCH-M1"
/tissue_type="liver"
184..1596
/note="secretory protein"
/codon_start=1
/product="aminopeptidase"
/protein_id="A043215.1"

CDS
1030 GTCTGCTGAGTGGACATTTGGACAGCTGGGATGTTGGGAGGGTGCATGATGATGC 1089

```

/db_xref="GI:5442032"
/translation="MRSLEFLIVHLALGSKAVPKNGVRSORTPREIKELIANVEDY
AKAIIALAVGYKXNSYERILGLIDVDPRLSGSKNYELAIOMONLODDLEMYH
LEQVRIPIHMERGESAVMLEPRIRHKAIIIGLSSISITPREGITAEVLVVASPELQR
ASEBARKIIVYNOPPYTGEKTVQVRQGANEAARKVAVASLIOSVASFISPTGIIQ
KODGVPKIPTACITVEDBEMSRMSRNGKIVIHDEMKAKTYPDSDSFVTEITGS
MYEEVAVLSGHDSDVQGLADDDGGAFLSWEALSLVLDLRPRTLRLVLAETAE
BOGIGASOYVELHKANISKYSLSMEADSGTFLPTGIQFTGSDKARAIKMEVNLLOP
LNTYKPSNSEGTDINPMIOAGVPGLSLDDLYKIFPFHSHSDTYTMDPKOMNVA
AVAAVAVVAVMDKDEMLPRS"
BASE COUNT 415 a 357 c 421 g 403 t
ORIGIN
Query Match 80.7%; Score 1084.8; DB 10; Length 1596;
Best Local Similarity 87.9%; Post. No. 2.4e-297;
Matches 1182; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
QY 1 AAGATGGCATCTCTTAAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGA 60
DB 250 AAGATGGTGTCTCTAGGACATTTTGGAAATATAAGAAATAGCCATTTATGAA 309
QY 61 GATGTTGCTAAAGCATCATCACTAGCTGTTTATGTTAAAGCCAGAAACAGATCTAT 120
DB 310 GATGTTGCTAAAGCATCATCACTAGCTGTTTATGTTAAATACCAAGAACCGTCTAT 369
QY 121 GAGCGATTGGCCTTCTGTTGATPACTGTTGACCCAGACTGAGTGGCTCCAGAACTTA 180
DB 370 GAGCGTTTGGACTTCTAGTTGATPACTGTTGACCCAGACTGAGTGGCTTAAAGAACTTA 429
QY 181 GAAAGAGCATCCAAATTTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTTAC 240
DB 430 GAGAAAGCTATTCAAATCATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTTAC 489
QY 241 CTGAGCCAGTGAAGATATCCCACTGGGAGAGGGAGAGAAATGCTGTGATGCTGAG 300
DB 490 CTGAGCCAGTGAAGATATCCCACTGGGAGAGGGAGAGAAATGCTGTGATGCTGAG 549
QY 301 CCAAGATTCATTAAGATAGGCAATCCGCGGCTTGGCAGCAGCAATTGGGACTCCGCCAGAA 360
DB 550 CTTGAAATTCACAGAGTGGCTATTTAGTTGTTGGCAGCAGCAATTGGGACTCCGCCAGAA 609
QY 361 GGCATTACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAATGTCAGAGAGGCGCTCA 420
DB 610 GGCATTACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAATGTCAGAGAGGCGCTCA 669
QY 421 GAAGCAAGAGGAGATGTTGTTTATTAACCAACTTACATCAACTCAAGAGCGGTG 480
DB 670 GAAGCAAGAGGAGATGTTGTTTATTAACCAAGCTTACACTGGCTATGAGAAAGACTGTG 729
QY 481 CAATACCGAAGCAGAGGGGGGGGAGAGAGCTGCCAAGGTGGGGCTTGGCATCTCATAT 540
DB 730 CAGTACCGGAGTGCAGGGAGCTGTGAGAGCTGCCAAGGTGGAGCTGTGGCATCTCATATC 789
QY 541 CGATCCGTGCGCTCTCTCTCATCTACAGTCTCAACAGAGTATTCAGAAATACAGAGAT 600
DB 790 CAATCAGTACGTTCTCTTTCATCTACAGTCTCAACAGAGTATTCAGAAATACAGAGAT 849
QY 601 GGCCTGCCAAATTCACAGCCTGTATTACGGTGAAGATGCAAGATGATGTCAAGA 660
DB 850 GGTGTGCCAAAGATTCACAGCCTGTATACAGTATGAAGATGAGAAATGATGTCTCA 909
QY 661 ATGCTTCTCATGGAGTCAAAATTTGTCACTACAGTAAAGATGGGGGCAAGACCTTACCA 720
DB 910 ATGCTTCTCTGTGGAAACAAATTTGTCACTTCTGAAATGGAGGAAAGACCTTATCCA 969
QY 721 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATTCAGAAACAGTT 780
DB 970 GATACGATTCCTTCAACACTGTGTGAGAGATCACTGGGAGCAATGTATTCAGAGAGAGTT 1029
QY 781 GTACTGTCAGTGAACATCTGACACGCTGGATGTTGGGAGGGTGCATGATGATGC 840
DB 1030 GTCTGCTGAGTGGACATTTGGACAGCTGGGATGTTGGGAGGGTGCATGATGATGC 1089

```

Qy 841 GGTGAGACCTTTATATCATGGAGACACTCTCATTTATTAAGATCTTGGCTGGCTCA 900
 |||||
 Db 1090 GGTGAGACCTTCATATCATGGAGACACTCTCATTTATTAAGATCTTGGCTGGCTCA 1149
 Qy 901 AAGAGACTCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 960
 |||||
 Db 1150 AAGAGACTCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1209
 Qy 961 CAGTATTCAGATTACAGAGGTAATATTTCCACTGCTGGCTGGCTGGCTGGCTGGCT 1020
 |||||
 Db 1210 CAGTATTCAGATTACAGAGGTAATATTTCCACTGCTGGCTGGCTGGCTGGCTGGCT 1269
 Qy 1021 GCAGAACCTTCTTACCACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1080
 |||||
 Db 1270 TCAGGAACCTTCTTACCACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1329
 Qy 1081 ATGAGAGGTTATGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1140
 |||||
 Db 1330 ATGAGAGGTTATGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1389
 Qy 1141 GAAGGACAGACATCACTTTTGGATCGAAGCTGGAGTGGCTGGAGCTGCTGAT 1200
 |||||
 Db 1390 GAAGGACAGACATCACTTTTGGATCGAAGCTGGAGTGGCTGGAGCTGCTGAT 1449
 Qy 1201 GACTTATCAAGATATTTCTTCCATCATCTCCACGAGACACCATGATGTCATGAT 1260
 |||||
 Db 1450 GACTTATCAAGATATTTCTTCCATCATCTCCACGAGACACCATGATGTCATGAT 1509
 Qy 1261 CCAAGACGATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 |||||
 Db 1510 CCAAGACGATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1569
 Qy 1321 ATGAGAGGTTATGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1344
 |||||
 Db 1570 ATGAGAGGTTATGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1593

RESULT 11
 AF131077 1726 bp mRNA linear ROD 03-JAN-2003
 LOCUS Rattus norvegicus liver annexin-like protein (LAL) mRNA, complete
 DEFINITION cds.
 ACCESSION AF131077 GI:7108712
 VERSION AF131077.1
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus (Norway rat); Chordata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1726)
 AUTHORS Della Fazio, M.A., Piobbico, D., Bartoli, D., Castellini, M.,
 Brancorsini, S., Viola Magni, M. and Servillo, G.
 TITLE lal-1: a differentially expressed novel gene during proliferation
 in liver regeneration and in hepatoma cells
 JOURNAL Genes Cells 7 (11), 1183-1190 (2002)
 MEDLINE 22378398
 PUBMED 12390252
 REFERENCES 2 (bases 1 to 1726)
 AUTHORS Servillo, G., Della Fazio, M.A., Piobbico, D., Bartoli, D.,
 Castellini, M., Brancorsini, S. and Viola Magni, M.
 TITLE Direct Submision
 JOURNAL Submitted (25-FEB-1999) Institute of General Pathology, University
 of Perugia, Policlinico Monteluce, Perugia 06100, Italy
 FEATURES
 source
 1..1726
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..1726
 /gene="LAL"
 63..1481
 /gene="LAL"

/note="Involved in liver regeneration; contains unique
 annexin domain at the N-terminus"
 /codon_start=1
 /product="liver annexin-like protein"
 /protein_id="AA036518.1"
 /db_xref="GI:7108713"
 /translation="MRFLPFLFVAVHLFSLGSGKAIYKSGVSRPFOEIKIEIANE
 DVAKAIINLAVGKQVRSYERGLVDVTGPRLSGKNLEKAIQIWMYOMLODGLN
 VHLIYVRIPIHMERGESAVWVPRIRHIALILGIGSGIGTPEGITLAVLAVSVELQ
 RASERAGKIVVYNOPPTDGYKQVQRERAVDAKYGAAVSLRSASRSYSPHG
 HQGVDGVPKIPKACITTEDKMSRFRSGKDIYHLKRGATYEPDTSFNTVAEIT
 GSKYPERVVLVSHLSDVDGQALDDGGAFSLVSWELSLVKDLRPRKTLRLVLT
 AERGGVGAQYVELEHFNLSKSLVMEADSGTFLPLPGLOFTGSDKARALMEKVMIL
 OPTNITKTFNDAGTDINFWIIOAGVPSASLRDLKYKFFRHHSGDTMTMDPKQNV
 AAIVMAVVAIVVADMEMLPRS"

BASE COUNT 474 a 385 c 428 g 439 t
 ORIGIN

Query Match 78.9%; Score 1060.8; DB 10; Length 1726;
 Best Local Similarity 86.8%; Pred. No. 1.6e-230;
 Matches 1167; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 1 AAGAGTGCATCTCTAAGAGACTTTGAAGAAATTAAGAGAAATAGCCAGCTGTGA 60
 |||||
 Db 135 AAGAGTGTGTTCTCGACCAACATTTCAAGAAATTAAGAGAAATAGCCACTATGAA 194
 Qy 61 GATGTGCTAAGCAATCATCAACTGCTGTTATGATGAAGCCGAGAACAGATCTAT 120
 |||||
 Db 195 GATGTGCTAAGCAATCATCAACTGCTGTTATGATGAAGAAATAGCCAGCTGTAT 254
 Qy 121 GAGCGATTGGCACTTCTGTTGATCTGTTGACCCAGACTGATGCTCCAGAACCTTA 180
 |||||
 Db 255 GAGCGTTGGGACTTCTAGTTGATAGTGTGACCCAGACTGATGCTCTAAGAACTTA 314
 Qy 181 GAAGAGCCATCCAAATTAATGACCAAACTGACGAGAATGGGCTGGAGAAAGTTAC 240
 |||||
 Db 315 GAGAAAGCTATCCAAATCATGTACCAAACTGACGAGAATGGGCTGGAGAAAGTCCAC 374
 Qy 241 CTGAGAGCAGTGAAGAAATACCCACTGGAGAGGGGAGAAAGATAGCTGATGCTGGAG 300
 |||||
 Db 375 CTGAGAGCAGTGAAGAAATACCCACTGGAGAGGGGAGAAAGATAGCTGATGCTGGAG 434
 Qy 301 CCAAGATTCATTAAGATAGCCATCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGAA 360
 |||||
 Db 435 CCGGAATTCACAAAGTTGGGATTTTGGGCTTTGGGCGGAGCATTTGGGACTCTCCAGAA 494
 Qy 361 GGCATTACAGCAGAGATTCTGGTGTGACCTCTTTGATGATGACGACAGAAAGGCTCA 420
 |||||
 Db 495 GGTATCACAGCAGAGAGTACTGGTGTGCTCTTTGTTGAACCTTCAAGAGGGGCAATCA 554
 Qy 421 GAAGCAAGAGGGAAGATTGTTGTTTATTAACCAACTTACATCACTCAAGAGCGGTG 480
 |||||
 Db 555 GAGGACAGAGGGAAGATTGTTGTTTATTAACCAAGCTTACACTGATGAGGAAACCTGTG 614
 Qy 481 CAATACCGAAGCAGAGGGGCGGTGGAAGCTGCCAAGGTGGGGGCTTTGGCAATCTCAT 540
 |||||
 Db 615 CAGTACCGGAGAGCGCGGAGGTGTGGAAGCTGCCAAGGTGGGGGCGGTGATCCCTCATC 674
 Qy 541 CGATCGGTGGCTCTCTTCTCATCTTACAGTCTCACAAGGATTTTCAGAAATACAGAGAT 600
 |||||
 Db 675 CGATCAGTAGCTTCTTTTTCATCTACAGTCTCACAAGGATTTTCAGAAATACAGAGAT 734
 Qy 601 GGCCTGGCCAAATTTCCACAGCCTGATTAACGCTGGAAGATGAGAAATGATGTCAAGA 660
 |||||
 Db 735 GGTGTGCCCAAGATTTCCACAGCCTGATTAACCAATTAAGATGAGAAATGATGTCTCGA 794
 Qy 661 ATGCTTCTCATGGGATCAAAATTTGATTCAGCTTAAGATGGGGGCAAAAGACTATCCCA 720
 |||||
 Db 795 ACGGCTCTCGTGGGAGCAAAATTTGATTCATCTGAAATGGAGCAAAAGACTATCCA 854
 Qy 721 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATTCAGAAACAGATT 780
 |||||
 Db 855 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATTCAGAAAGATT 914

QY	781	GTACTGGCATGTGACATCTCTGACAGCTGGGATGTTGGGACAGGGGTCACATGAATGATGGC	840
Db	915	GTCTCTGGTCACTGGACATCTTGGACAGCTGGGACGTCTGGGACGGGTGCACCTGAATGAATGGC	974
QY	841	GGTGGAGCTTTTATATATCAATGGGAAGCACTCACTTATTTAAAGATCTTTGGGCTCGTCCA	900
Db	975	GGTGGAGCTTTCAATATCAATGGGAAGCACTCACTTGTTTAAAGATCTTTGGGCTCGTCCA	1034
QY	901	AAGAGCATCTTCGGGCTGGTGTCTGTGGACCTCAGAGAAGACAAAGTGGAGTTGGTGCCTTC	960
Db	1035	AAGAGCATCTTCGGGCTGGTGTCTGTGGACCCGACAGAAACAAAGAGGGGTTGGTGCCTTC	1094
QY	961	CAGATTTATCAGTTACACAAAGTAATATTTTCCAACTACACTCTGGTGAATGGAGTCTGAC	1020
Db	1095	CAGATTTATGAGCTACATTAAGGCANATATTTTCCAGTACATGTTGGTGAATGAAGGCTGAC	1154
QY	1021	GCAGGAACCTTCTTACCCACTGGGCTGCAATTCACCTGGCAGTGAAGAGGCCAGGAGCCATC	1080
Db	1155	TCAGGAACCTTCTTATCCACTGGGCTGAGTTCAACCGGCAATTGACAAAGGCCAGGGGCTATC	1214
QY	1081	ATGAGAGAGGTTATGAGGCTGTGCAAGCCCTCAATATCACTCAGGTCCTGAGCCATGGA	1140
Db	1215	ATGAGAGAGGTCATGAGGCTCTCTGCACCCCTCAATATCAACCAAGGTCCTTTAATGATGCA	1274
QY	1141	GAAAGGACAGACATCAACTTTTGGATCCAAACCTGGAAGGCTTGGAGCCAGTCACTTGAT	1200
Db	1275	GAAAGGAATGCAATTAACCTTCTTGATCCAAACCTGGAAGGCTTGGAGCCAGTCTCGAGAT	1334
QY	1201	GACTTATATCAAGTATTTCTTCTTCATCACTCCACGGAGACACATGACCTGTATGAT	1260
Db	1335	GACTTGTATCAAGTATTTCTTTTCCATCATTTCCATGSAAGACACATGACCTGCCATGAT	1394
QY	1261	CCAAAGCAGATGAATGTTGCTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTGCAGAC	1320
Db	1395	CCAAAGCAGATGAATGTTGCTGCTGCTGCTGTTGGGCTGTTGCTTACGTTAGTTGTGCGAGAC	1454
QY	1321	ATGGAAGAAATGCTGCCATGATCC	1344
Db	1455	ATGGAAGAAATGCTGCCAGGTTCC	1478

RESULT 12	AX400829	1778 bp	DNA	linear	PAT 06-JUN-2002
LOCUS	AX400829				
DEFINITION	Sequence 505 from Patent WO0210453.				
ACCESSION	AX400829				
VERSION	AX400829.1	GI:21337009			
KEYWORDS					
SOURCE					
ORGANISM	Rattus norvegicus (Norway rat)				
	Rattus norvegicus				
	Eulatyvola, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
	Rattus.				
REFERENCE	1.				
AUTHORS	Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and Blashoff, M.R.				
TITLE	Molecular toxicology modeling				
JOURNAL	Patent: WO 0210453-A 505 07-FEB-2002;				
	Gene Logic, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..1778				
	/organism="Rattus norvegicus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10116"				
	/note="EMBL/GenBank Accession No. AF097723"				
BASE COUNT	488 a	391 c	442 g	457 t	
ORIGIN					

Query Match	78.9%	Score 1060.8	DB 6	Length 1778
Best Local Similarity	86.8%	Pred. No. 1.6e-290		
Matches 1167, Conservative	0	Mismatches 177	Indels 0	Gaps 0

OY	1	AAAGATGGCACTCTTAAGAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGGGA	60
Db	193	AAGAGTGTGTTTCTCGCGAACTTTTCAAGAAATTAAGAGAAATAGCCACTATGAA	252
OY	61	GATTTGCTAAGCAATCATCAACTAGCTGTTTATGTAAGCCCAAGACATCTAT	120
Db	253	GATTTGCTAAGCAATTTATCAACTTGTCTGTTTATGGAATAATCCAGAACCGGTCTAT	312
OY	121	GAGCGATTGGCACTTCTGGTTGATCTGTTGGACCCGACTGAGTGGCTCCAGAACCTA	180
Db	313	GAGGTTTGGGACTTCTAGTTGATCTGTTGACCCGACGTGAGTGGCTCTAAGAACCTA	372
OY	181	GAAGAAAGCCATCCAAATTAATGTAACCAAACTGAGCAAGATGGGCTGGAGAAAGTTAC	240
Db	373	GAGAAAGCTATCCAAATCATGTACCAAAACCTGCAAGAAATGGGCTGGAGAAAGTCCAC	432
OY	241	CTGAGAGCAGTGAATAATCCCACTGGAGAGGGGAGAAATAATCAGCTGTGATCTGGAG	300
Db	433	CTGAGAGCAGTGAATAATCTCACTGGGAGGAGGGGCGAAGATTCGACGTGATGCTGTG	492
OY	301	CCAGAAATCATTAAGATAGCCATCTGGGTTCTGGACAGCATTTGGGACTCTCCAGAA	360
Db	493	CTCGAAATTCACAAAGTTGGCTATTTTATGGCTTGGCGGACAGCTATGGGACTCTCCGAA	552
OY	361	GGCAATTACAGAGAAAGTTCTGGTGGTGAACCTCTTTGATGAACTGGCAGAGAGGGCTCA	420
Db	553	GGATATCACAGAGAAAGTACTGGTGGTGGCCCTCTTTGTTGAACTTCAAGAGGGGATCA	612
OY	421	GAAGCAAGAGGAAAGATTTGTTGTTTATTAACAACCTTACATCACTACTCAAGAGCGGTG	480
Db	613	GAGGCAGAGGAGGAAGATTTGTTGTTTATTAACAAGCTTACACTGACTATGGGAAAACTGTG	672
OY	481	CAATACCGAAACGAGGGGGCGGTGGAAAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCAT	540
Db	673	CAGTACCGGAGCGCGAGCTGTGAGAGCTGCCAAGGTGGGGGCGGTGGCATCTCTCATC	732
OY	541	CGATCCGTGGCTCTCTCTCTCACTTAAGTCTTACACAGATTTACAGAAATACCAAGAT	600
Db	733	CGATCAAGTACGCTCTTTTCCATCTTACATGCTTACACAGGTCTACAGGATATCAAGAT	792
OY	601	GGCGTGGCCAAAAATTCACAAGCGCTGTATTAACGGTGGAGAGTGGAGAAATGATGTCAAGA	660
Db	793	GGTGTGCCCAAGATTCACAAGCTGTATTCACAATAGAAATGAGAAATGATGTCTCGA	852
OY	661	ATGGCTTCTCAATGGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGGCAAAAGCTTACCA	720
Db	853	ATGGCTTCTGTGGGGGCAAAATTTGTCAATTCAGCTAAAGATGGGAGCAAAAGCTTATCA	912
OY	721	GATATGATTTCTTCAACACTGTAGCAGAGATTCATGGGAGCAAAATTCAGAACAGGTT	780
Db	913	GATACAGATTTCTTCAACACTGTGTGACAGATTCATGGGAGCAAAATTCAGAGAAAGTT	972
OY	781	GTAATGGTACGTGACATCTGGACAGGTGGGAGTGTGGGACGGGTGCATGAGATGAGGC	840
Db	973	GTCCTGGTACGTGACATCTGGACAGGTGGGAGCTGTGGGACGGGTGTGATGAGATGAGC	1032
OY	841	GGTGGAGCTTTATATCATGGGAGCACTCTCACTTTTAAAGATCTTGGGCTGCGTCA	900
Db	1033	GGTGGAGCTTTCAATCATGGGAGCACTCTCACTTTTAAAGATCTTGGGCTGCGTCA	1092
OY	901	AAGAGACTCTGGCGGCTGTGTCTGGAATGACAGAGAACAGGTTGATGTGCTTC	960
Db	1093	AAGAGACTCTGGCGGCTGTGTCTGGAACGACAGAGAACAGGATTTGATGTGCTTC	1152
OY	961	CAGATTAATCATGTTACCAAGGTAAATTTCCAACTACAGTCTGGGTGATGGAATCTGAC	1020
Db	1153	CAGATTAATCATGTTACCAAGGTAAATTTCCAACTACAGTCTGGGTGATGGAATCTGAC	1212
OY	1021	GCAGAAACTTCTTACCAACTGGGCTCAATTCATCTGGCAGTGAAGAGCGCAGGCGCATC	1080
Db	1213	TCAGAACTTCTTACCAACTGGGCTCAATTCATCTGGCAGTGAAGAGCGCAGGCGCATC	1272
OY	1081	ATGAGAGAGGTTATGAGCTGCTGACGCCCTCAATATCACTCAGGTCTGAGGCATGGA	1140

Db 1273 ATGAGAGAGTATGATGCTTCCTGCAACCCCTCAATATACCAAGTCTTTATGATGCA 1332

Qy 1141 GAAGGACAGACATCACTTTGGATCCAGCTGAGTGCCTGGAGCAGTCTACTGAT 1200

Db 1333 GAAGGACAGACATCACTTTGGATCCAGCTGAGTGCCTGGAGCAGTCTACTGAT 1332

Qy 1201 GACTTATACAGATTTTCTTTCATCACTCCACGAGACACACATGATGTCATGAT 1260

Db 1393 GACTTATACAGATTTTCTTTCATCACTCCACGAGACACACATGATGTCATGAT 1452

Qy 1261 CCAAGGACAGATTTTCTTTCATCACTCCACGAGACACACATGATGTCATGAT 1330

Db 1453 CCAAGGACAGATTTTCTTTCATCACTCCACGAGACACACATGATGTCATGAT 1512

Qy 1321 ATGAGAGAGTATGATGCTTCCTGCAACCCCTCAATATACCAAGTCTTTATGATGCA 1344

Db 1513 ATGAGAGAGTATGATGCTTCCTGCAACCCCTCAATATACCAAGTCTTTATGATGCA 1536

RESULT 14

LOCUS AF009513 1716 bp mRNA linear ROD 01-JUN-1998

DEFINITION Mus musculus hematopoietic lineage switch 2 (HLS2) mRNA, complete cds.

ACCESSION AF009513 GI:3169728

VERSION AF009513.1

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1716) Williams, J.H., Chan, C.-Y. and Klinken, S.P. Hematopoietic lineage switch 2 (HLS2), a novel mRNA species induced during an erythroid to myeloid lineage switch Unpublished

JOURNAL 2 (bases 1 to 1716) Williams, J.H., Chan, C.-Y. and Klinken, S.P. Direct Submission

REFERENCE 1 (bases 1 to 1716) Submitted (20-JUN-1997) Biochemistry Department (Laboratory of Cancer Medicine), University of Western Australia, Royal Perth Hospital, Wellington Street, Perth, WA 6001, Australia

JOURNAL 1. 1716 Location/Qualifiers

FEATURES

source

1. 1716 /organism="Mus musculus"

gene /mol_type="mRNA"

1. 1716 /db_xref="taxon:10090"

CDS /gene="HLS2"

97..1398 /gene="HLS2"

/codon_start=1

/product="hematopoietic lineage switch 2"

/protein_id="AACI7945.1"

/db_xref="GI:3169728"

/translation="MRSFLPLFVILHLALSGKAVFGNGVRSORTFRIKEIANYEDV AKVILAVYGVKONRSYERLGLIVDVPRLSGSKLKKLAKIQIMYONLOODGLFNH LEQVRIPIHMERGESAVMLBRIRHKAKILGLGSGITGRPGGITHFVIVASFDLQNR ASEARGLIIVNDPYTGYEKTQYRVGAAVAAKVAASLQISVASFISPTGIGQ KYDGVKPIPTACTIVDAEMWSMARGNKIVTHLMGAKTYPDTSFNTVAITGS MYEDVQGLADLDGGAFLISWEALSLVKDGLRKRTRLRLVLTAAEEGGIGASQYSE LHKANISKYSVLMEADSGTFLPTGLDFTGSDPKARAIKKEYWNLLQPLNVTIKVPSNGSG TDINFWIQAVPGASLRDDLYKYPFRPHSHSDTMTSIQSR"

BASE COUNT 476 a 370 c 414 g 456 t

ORIGIN

Query Match 74.4%; Score 1000.4; DB 10; Length 1716;

Best Local Similarity 85.6%; Pred. No. 2,66-273;

Matches 1151; Conservative 0; Mismatches 156; Indels 37; Gaps 2;

Qy 1 AAGATGCACTCTTAAAGAGACTTTGAAGAATAAAGAAAGATAGCCAGCTGTGGA 60

Db 163 AAGATGCTGTTCTCAGCGAATCTTCAGAGAAATTAAGAAAGATAGCCAGCTATGAA 222

Qy 61 GATGTGCTAAAGCAATCATCAACCTAGCTGTTATGTTAAAGCCAGAACAGATCTAT 120

Db 223 GATGTGCTAAAGCAATCATCAACCTGCTGTTATGTTAAAGCCAGAACAGATCTAT 282

Qy 121 GAGCGATTGGCACTTCTGTTGATPACTGTTGACCCAGACTGATGGCTCCAGAACCTA 180

Db 283 GAGCGTTGGGACTTCTGTTGATPACTGTTGACCCAGACTGATGGCTCCAGAACCTA 342

Qy 181 GAAAAAGCATCCAAATTATGTCACAAACCTGTCAGCAAGATGGGCTGAGAAAGTTTAC 240

Db 343 GAGAAAGCTATTCAAATCATGTCACAAACCTGTCAGCAAGATGGGCTGAGAAAGTTTAC 402

Qy 241 CTGAGCGCAATGAGAAATACCCCACTGGGAGAGGGGAGAAATACAGCTGATCTGAG 300

Db 403 CTGAGCGCAATGAGAAATACCCCACTGGGAGAGGGGAGAAATACAGCTGATCTGAG 462

Qy 301 CCAAGAAATTCATTAAGATAGCCATCTGGGATCTTGGACAGCATTTGGGACTCTCCAGAA 360

Db 463 CTGCAATTCACAAGATAGCCATCTGGGATCTTGGACAGCATTTGGGACTCTCCAGAA 522

Qy 361 GGCATTCACAGAAAGTTCTGTTGTCGACCTCTTTGATGATGATGTCAGAAAGGCTCA 420

Db 523 GGCATTCACAGAAAGTTCTGTTGTCGACCTCTTTGATGATGATGTCAGAAAGGCTCA 582

Qy 421 GAAGCAAGAGGAAAGATTTGTTTATACCAACCTTACATCACTCAAGGACGGTG 480

Db 583 GAAGCAAGAGGAAAGATTTGTTTATACCAACCTTACATCACTCAAGGACGGTG 642

Qy 481 CAATACCGAAGCAAGGAGGAGGAGTGAAGCTGCCAAGGTGGGGCTTTGGATCTTCATT 540

Db 643 CAATACCGGAGTGAAGGAGGAGTGAAGCTGCCAAGGTGGGGCTTTGGATCTTCATT 702

Qy 541 CGATCGTGCCCTCTCTTCTCATCTAGTACAGTCTTCACACAGATTTCCAGAAATACAGAT 600

Db 703 CAATACGATACCTCTCTTCTCATCTAGTACAGTCTTCACACAGATTTCCAGAAATACAGAT 762

Qy 601 GGGGTGGCCAAATTCACACAGGCTGATTAAGCGTGAAGATGAGAAATGATGTCAGA 660

Db 763 GGGGTGGCCAAATTCACACAGGCTGATTAAGCGTGAAGATGAGAAATGATGTCAGA 822

Qy 661 ATGGCTTCTCATGGGATCAAAATTTGTCATTCACTAAGTAAAGATGGGGCAAGACCTACCA 720

Db 823 ATGGCTTCTCATGGGATCAAAATTTGTCATTCACTAAGTAAAGATGGGGCAAGACCTACCA 882

Qy 721 GATTCGATTTCTTCAACATGTAAGCAGATCACTGGGAGCAATATCCAGAACAGGTT 780

Db 883 GATTCGATTTCTTCAACATGTAAGCAGATCACTGGGAGCAATATCCAGAACAGGTT 935

Qy 781 GTACTGTGATGATGATCTGACACAGCTGGAGATTTGGGAGGTCATGATGATGCG 840

Db 936 -----GAAATGGGAGGAGGTCATGATGATGCGT 966

Qy 841 GGTGAGACCTTTATATCATGAGGAGACCTCTCACTTTAAAGATCTTGGGCTGCTGCA 900

Db 967 GGTGAGACCTTTATATCATGAGGAGACCTCTCACTTTAAAGATCTTGGGCTGCTGCA 1026

Qy 901 AAGAGCACTTGGGCTGCTGCTGCTGCACTGCAAGAGACAGAGGAAATTTGGTCTCC 960

Db 1027 AAGAGCACTTGGGCTGCTGCTGCTGCACTGCAAGAGACAGAGGAAATTTGGTCTCC 1086

Qy 961 CAGATTATAGTTACCAAGGATTAATTTCCAACTCAAGTCTGATGATGAGTCTGAC 1020

Db 1087 CAGATTATAGTTACCAAGGATTAATTTCCAACTCAAGTCTGATGATGAGTCTGAC 1146

Qy 1021 GCAAGAACTTTTATACCACTGGGCTGCAATTCATGTCAGTGAAGAAAGCCAGGCGCATC 1080

Db 1147 TCAGGAACCTTTTATACCACTGGGCTGCAATTCATGTCAGTGAAGAAAGCCAGGCGCATC 1206

Qy 1081 ATGAGAGAGTTATAGCTGCTGCTGCAAGCCCTCAATATCACTAGTCTGAGGACATGGA 1140

Db 1207 ATGAGAGAGTTATAGCTGCTGCTGCAAGCCCTCAATATCACTAGTCTTATGATGGA 1266

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 12:51:18 ; Search time 323.93 Seconds

(without alignments)
11200.072 Million cell updates/sec

Title: US-09-745-763-35_COPY_171_1514

Perfect score: 1344

Sequence: 1 AAGAAATGCATCTTAAGAG.....AAGAAATGCTGCTAGCTCC 1344

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2004.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	1851	19	AAV82779
2	1344	100.0	1851	24	ABO92016
3	1340.8	99.8	1767	21	AAA40493
4	1340.8	99.8	1778	18	AAV02296
5	1340.8	99.8	1863	21	AAZ98034
6	1340.8	99.8	1863	22	AAD11647
7	1340.8	99.8	1863	24	ABK69743
8	1340.8	99.8	1863	25	ACC50817

Result No.	Score	Query Match	Length	DB ID	Description
9	1340.8	99.8	1863	25	ABZ71453
10	1340.8	99.8	1884	21	AAZ58313
11	1340.8	99.8	1923	21	AAZ98139
12	1340.8	99.8	2077	25	ACC50603
13	1340.8	99.8	2077	25	ABZ71331
14	1336	99.4	1860	22	AAK94491
15	1283.6	95.5	1895	22	AAH99703
16	1224.8	91.1	1784	23	AAZ73592
17	1060.8	78.9	1778	24	ABK62598
18	772.4	57.5	1134	21	AAZ98101
19	772.4	57.5	1134	22	AAZ11714
20	772.4	57.5	1134	24	ABK69810
21	772.4	57.5	1134	25	ACC50818
22	772.4	57.5	1134	25	ABZ71454
23	770.8	57.4	895	21	AAA44369
24	448.6	33.4	609	20	AAV87487
25	398.2	29.6	642	22	AAK92181
26	398.2	29.6	642	22	AAK93447
27	212.8	15.8	314	20	AAV86369
28	204.6	15.2	462	18	AAZ97398
29	204.6	15.2	462	18	AAZ98060
30	204.6	15.2	462	18	AAV02319
31	170.4	12.7	441	22	AAK56831
32	140.6	10.5	317	20	AAZ40587
33	114.2	8.5	492	22	AAK92932
34	90.8	6.8	394	14	AAO61260
35	61.4	4.6	424	23	AAZ53591
36	60.2	4.5	65	24	ABN53183
37	60	4.5	60	24	ABN56212
38	42	3.1	784	22	AAH05499
39	42	3.1	2200	22	AAH15906
40	42	3.1	2200	22	AAH78065
41	42	3.1	2408	22	AAH99550
42	41	3.1	65	24	ABN28454
43	39.4	2.9	975	21	AAZ77173
44	39.4	2.9	1556	24	AAZ33726
45	39.4	2.9	2319	25	ABX63579

ALIGNMENTS

RESULT 1
ID AAV82779 standard; CDNA; 1851 BP.
XX AAV82779;
XX 25-FEB-1999 (first entry)
DB Clone bu45_2 isolated from human adult placenta CDNA library.
XX Secreted protein; nutritional activity; immune stimulating; vaccine;
XX * suppressing activity; haematopoiesis regulating activity;
XX tissue growth activity; activin; inhibin activity; chemotaxis;
XX chemokine activity; haemostasis; thrombolytic activity; receptor;
XX ligand; anti-inflammatory; cadherin; tumour invasion suppressor;
XX tumour inhibition; gene therapy; ds.
XX Homo sapiens.
XX OS
XX PN WO9842739-A2.
XX PD 01-OCT-1998.
XX PF 20-MAR-1998; 98WO-US05653.
XX PR 19-MAR-1998; 98US-0044466.
XX PR 21-MAR-1997; 97US-0822167.
XX (GENY) GENETICS INST INC.
XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

US2002065394-A1.
30-MAY-2002.
22-DEC-2000; 2000US-0745763.
18-MAR-1998; 98US-0040963.
(JACO/) JACOBS K.
(MCCO/) MCCOY J M.
(LAVA/) LAVALLIE E R.
(COLL/) COLLINS-RACIE L A.
(EVAN/) EVANS C.
(MERB/) MERBERG D.
(TREA/) TREACY M.
(SPAU/) SPAULDING V.
Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C,
Merberg D, Treacy M, Spaulding V;
WPI; 2002-582343/62.
P-PSDB; ABP61800.
Novel secreted or transmembrane protein and polynucleotide encoding the
protein, useful for diagnosis and treatment of neurological disorders,
cancer, autoimmune diseases, bone disorders and lung or liver fibrosis
Claim 50; Page 113-114; 284pp; English.
The invention relates to human secreted or transmembrane protein (I),
their fragments and is encoded by specific complementary deoxyribonucleic
acid (CDNA) inserts (II), where the protein is substantially free from
other mammalian proteins. (I) are useful for preventing, treating or
ameliorating a medical condition, especially immunological treatment or
prevention of tumours. (I) exhibits activity relating to angiogenesis,
cytokine, cell proliferation, cell differentiation, anti-inflammatory,
stem cell growth factor activity and activin or inhibin-related
activities. (I) can be used to manipulate stem cells in culture to give
rise to neuroepithelial cells that can be used to augment or replace
cells damaged by illness, autoimmune disease, accidental damage or
genetic disorders. (I) induces the proliferation of neural cells and
regeneration of nerve and brain tissue and is useful for the treatment of
central and peripheral nervous system diseases and neuropathies, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis. (I) is involved in chemotactic or chemokinetic
activity, regulation of haematopoiesis and is useful for treating myeloid
or lymphoid cell disorders, platelet disorders such as thrombocytopenia
and for regeneration of bone, cartilage, tendon, ligament and/or nerve
tissue growth and in tissue repair, healing of burns, incisions, ulcers,
for treating osteoporosis, osteoarthritis, bone degenerative disorders or
periodontal disease. (I) is also useful for gut protection or
regeneration and treatment of lung or liver fibrosis, reperfusion injury
in various tissues, various immune deficiencies and disorders including
severe combined immunodeficiency (SCID), bacterial or fungal infections,
autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
such as asthma or other respiratory problems. (II) is useful to express
recombinant protein, as markers for tissues in which the corresponding
protein is preferentially expressed and in gene therapy. The present
sequence is that of a polynucleotide of the invention.
Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;
Query Match 100.0%; Score 1344; DB 24; Length 1851;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AAGATGCGATCTCTAAGAGAGCTTTGAAGAAATTAAGAAATATGCGAGCTGTGCA 60
171 AAGATGCGATCTCTAAGAGAGCTTTGAAGAAATTAAGAAATATGCGAGCTGTGCA 230
61 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTGTAAGCCAGACAGATCTTAT 120

231 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTGTAAGCCAGACAGATCTTAT 290
121 GAGCGATTGGGCACTTCGTTGATGATCTGTGACCCGACGATGGCTCCAGAACCTTA 180
291 GAGCGATTGGGCACTTCGTTGATGATCTGTGACCCGACGATGGCTCCAGAACCTTA 350
181 GAAAGAGCCATCCAAATTAATGTAACCAAACTGACGAGAGATGGCTGAGAAAGTTTAC 240
351 GAAAGAGCCATCCAAATTAATGTAACCAAACTGACGAGAGATGGCTGAGAAAGTTTAC 410
241 CTGGAGCCAGTGAATATCCCACTGGAGAGAGGAGAGAAATCACTGTGATGCTGAG 300
411 CTGGAGCCAGTGAATATCCCACTGGAGAGAGGAGAGAAATCACTGTGATGCTGAG 470
301 CCAAGATTCATTAAGATAGGCAATCTGGGATCTGGGAGAGCAATTTGGGATCTCCCGAA 360
471 CCAAGATTCATTAAGATAGGCAATCTGGGATCTGGGAGAGCAATTTGGGATCTCCCGAA 530
361 GGCATTACAGCAGAGATTTGTTGATGATCTGTTGATGATCTGTTGATGATCTGTTGAT 420
531 GGCATTACAGCAGAGATTTGTTGATGATCTGTTGATGATCTGTTGATGATCTGTTGAT 590
421 GAAGCAAGAGGAGAGATTTGTTGATGATCTGTTGATGATCTGTTGATGATCTGTTGAT 480
591 GAAGCAAGAGGAGAGATTTGTTGATGATCTGTTGATGATCTGTTGATGATCTGTTGAT 650
481 CAATACCGAAGCAGAGGAGGAGGAGAGAGCTGCAAGAGTGGGGCTTGGATCTTCAT 540
651 CAATACCGAAGCAGAGGAGGAGGAGAGAGCTGCAAGAGTGGGGCTTGGATCTTCAT 710
541 CGATCCGTGGGCTCTCTCTCATCTACATGATCACTGATCACTGATCACTGATCACTGAT 600
711 CGATCCGTGGGCTCTCTCTCATCTACATGATCACTGATCACTGATCACTGATCACTGAT 770
601 GGGCTGCCCCAAATTTCCAAACAGCTGATTAACGGTGAAGATGAGATGATCAAGA 660
771 GGGCTGCCCCAAATTTCCAAACAGCTGATTAACGGTGAAGATGAGATGATCAAGA 830
661 ATGGCTTCTCATGGAGATCAAAATTTGATTCAGTAAAGATGGGGGCAAAAGACTTACCA 720
831 ATGGCTTCTCATGGAGATCAAAATTTGATTCAGTAAAGATGGGGGCAAAAGACTTACCA 890
721 GATCTGATTCCTTCAACACTGTGACAGATCACTGAGAGCAATATTCAGAACAGTT 780
891 GATCTGATTCCTTCAACACTGTGACAGATCACTGAGAGCAATATTCAGAACAGTT 950
781 GTACTGTCACTGATCACTGTGACAGATCACTGAGAGCAATATTCAGAACAGTT 840
951 GTACTGTCACTGATCACTGTGACAGATCACTGAGAGCAATATTCAGAACAGTT 1010
841 GGTGAGACCTTTATATATGAGGAGAGCACTCTCACTTTTAAAGATCTTGGGCTGCTTCA 900
1011 GGTGAGACCTTTATATATGAGGAGAGCACTCTCACTTTTAAAGATCTTGGGCTGCTTCA 1070
901 AAGAGAGCTGCGGGCTGAGTCTGAGTCACTGCAAGAAACAAAGGTGAGTTGAGCTTC 960
1071 AAGAGAGCTGCGGGCTGAGTCTGAGTCACTGCAAGAAACAAAGGTGAGTTGAGCTTC 1130
961 CAGTATTATCAGTATCAACAAGATTAATTTTCAACTGATGATGATGATGATGATGATGAT 1020
1131 CAGTATTATCAGTATCAACAAGATTAATTTTCAACTGATGATGATGATGATGATGATGAT 1190
1021 GCAGAGAACCTTTTATCCCACTGGGCTGATTCATCTGCGAGTAAAGGCCAGGGCCATC 1080
1191 GCAGAGAACCTTTTATCCCACTGGGCTGATTCATCTGCGAGTAAAGGCCAGGGCCATC 1250
1081 ATGAGAGAGGTTATGAGCTGCTGAGAGCCCTCAATTCATCTGAGGCTGAGGAGATGAG 1140
1251 ATGAGAGAGGTTATGAGCTGCTGAGAGCCCTCAATTCATCTGAGGCTGAGGAGATGAG 1310
1141 GAAAGGACAGACATCACTTTTGGATCCAACTGAGAGCTGAGGAGGAGGATCTTATGAT 1200

Db	1311	GAAGGAGACAGACATCAACTTTTGGATCCAAAGTCGGAGTCCCTGGAGCCAGCTACTGAT	13
Qy	1201	GACTTATACAGTATTTCTTTCCATCACTCCCAAGAGACACATGACTGTATGAT	126
Db	1371	GACTTATACAGTATTTCTTTCCATCACTCCCAAGAGACACATGACTGTATGAT	143
Qy	1261	CCAAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTGTTCTTATGTTGTGCAGC	132
Db	1431	CCAAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTGTTCTTATGTTGTGCAGC	149
Qy	1321	ATGAAGAATGCTGCTAGTCC	1344
Db	1491	ATGAAGAATGCTGCTAGTCC	1514
RESULT 3			
AAA0493	standard; cDNA; 1767 BP.		
AAA0493			
DT	16-NOV-2000 (first entry)		
DE	Human fetal kidney cDNA fragment AM282_11.		
KM	Secreted protein; cytosolic; immunostimulatory; antimicrobial;		
KM	antiviral; immunosuppressive; antiinflammatory; vulnerrary; cytokine;		
KM	cell proliferation; differentiation; regulator; treatment; tumor;		
KM	autoimmune disease; inflammatory disorder; wound; microbial infection;		
KM	viral disease; graft versus host reaction suppression; ss.		
OS	Homo sapiens.		
XX	WO200037630-A1.		
XX	29-JUN-2000.		
PD	22-DEC-1999; 99MO-US31005.		
XX	23-DEC-1998; 98US-0220876.		
XX	(GENY) GENETICS INST INC.		
PA	Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;		
PI	Merberg D, Treacy M, Bowman MR;		
DR	P-PsDB; AAB10229.		
XX	WPI; 2000-442661/38.		
PT	Secreted human proteins AS296-11 and AS34-11, useful for treating		
PT	tumors, autoimmune diseases, inflammatory disorders, wounds, microbial		
PT	infections and viral diseases -		
PS	Disclosure; Page 198; 293pp; English.		
XX	This invention describes novel secreted human proteins (I) which have		
CC	cytostatic, immunostimulatory, antimicrobial, antiviral,		
CC	immunosuppressive, antiinflammatory and vulnerary activity and which act		
CC	as cytokine, cell proliferation or differentiation regulators. (I)		
CC	is useful for treating tumors, autoimmune diseases, inflammatory		
CC	disorders, wounds, microbial infections and viral diseases. (I) is also		
CC	useful for suppressing graft versus host reaction. AAA0490-A40580		
CC	represent cDNA fragments that encode the secreted proteins		
CC	AAB10226-B10288 described in the method of the invention.		
XX	Sequence 1767 BP; 512 A; 384 C; 418 G; 453 T; 0 other;		
Qy	Query Match 99.8%; Score 1340.8; DB 21; Length 1767;		
	Best Local Similarity 99.9%; Pred. No. 0;		
	Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
	1 AAGATGCACTTCCTTAAGAGCACTTTGAAAGAAATTAAGAAAGAAATGCTGCTGGA 60		

D	78	AAGATGGCACTCTGTAAGAGGACTTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGGA	137
Q	61	GATGTGCTTAAGCAATCATCACTAGCTGTTATGTGTAAGCCAGAACGATCTTAT	120
D	138	GATGTGCTTAAGCAATCATCACTAGCTGTTATGTGTAAGCCAGAACGATCTTAT	197
Q	121	GAGGCAATTGGCACTTCGTGGTGTGATACGTGTGGAACCCAGACGAGTGGCTCCAGAAACCTA	180
D	198	GAGGCAATTGGCACTTCGTGTGTGATACGTGTGGAACCCAGACGAGTGGCTCCAGAAACCTA	257
Q	181	GAAGAAAGCCATCCAAATTTATGTATACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTTAC	240
D	258	GAAGAAAGCCATCCAAATTTATGTATACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTTAC	317
Q	241	CTGAGACCAAGTGAAGATACCCCACTGGGAGAGGGAGAAAGATCAGCTGTGTATGCTGGAG	300
D	318	CTGAGACCAAGTGAAGATACCCCACTGGGAGAGGGAGAAAGATCAGCTGTGTATGCTGGAG	377
Q	301	CCAAGAAATTAAGATAGATAGCATCTGGGATCTTGGACAGCATTTGGGACTCCTCCAGAA	360
D	378	CCAAGAAATTAAGATAGATAGCATCTGGGATCTTGGACAGCATTTGGGACTCCTCCAGAA	437
Q	361	GGCAATTAACAGCAGAAAGTTCTGGTGTGACCTCTTTCGATGAACTGCAGAGAAAGGGCTCA	420
D	438	GGCAATTAACAGCAGAAAGTTCTGGTGTGACCTCTTTCGATGAACTGCAGAGAAAGGGCTCA	497
Q	421	GAAGCAAGAGGGAAGATTTGTTTATATACCAACTTATCATCACTACAGAGCGGTG	480
D	498	GAAGCAAGAGGGAAGATTTGTTTATATACCAACTTATCATCACTACAGAGCGGTG	557
Q	481	CAATACCGGAAGGCAAGGGGGCGGTGGAAGCTGCCAAGGTGGGGCGCTTGGCATCTCTCAT	540
D	558	CAATACCGGAAGGCAAGGGGGCGGTGGAAGCTGCCAAGGTGGGGCGCTTGGCATCTCTCAT	617
Q	541	CGATCCGGGCGCTCCTTCTCATCTACAGTCTCAACAGGTATTTCAAGAAATACAGAGAT	600
D	618	CGATCCGGGCGCTCCTTCTCATCTACAGTCTCAACAGGTATTTCAAGAAATACAGAGAT	677
Q	601	GGCGGTGCCAAAAATTCCAACAGCCTGTATTAACGGTGAAGATGCAGAAATGATGTCAAGA	660
D	678	GGCGGTGCCAAAGATTCCAACAGCCTGTATTAACGGTGAAGATGCAGAAATGATGTCAAGA	737
Q	661	ATGGCTTCTCATGGGATTCAAAATTTGTCAATTCAGCTTAAAGATGGGGGGCAAAACCTACCCA	720
D	738	ATGGCTTCTCATGGGATTCAAAATTTGTCAATTCAGCTTAAAGATGGGGGGCAAAACCTACCCA	797
Q	721	GATACGTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATACCAAGACAGGTT	780
D	798	GATACGTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATACCAAGACAGGTT	857
Q	781	GTACTGTGCACTGAGACATCTTGACAGCTGGGATGTTGGCAGGGTCCATGATGATGGC	840
D	858	GTACTGTGCACTGAGACATCTTGACAGCTGGGATGTTGGCAGGGTCCATGATGATGGC	917
Q	841	GGTGAAGCTTTATATCATGGGAAGCACTCACTTATTAAGATCTTGGGCTGGGTCCA	900
D	918	GGTGAAGCTTTATATCATGGGAAGCACTCACTTATTAAGATCTTGGGCTGGGTCCA	977
Q	901	AAGAGGACCTCGGGGCTGGTGCCTGAGACCTGCAGAAAGAAAGAGGTGGAGTGGCTTTC	960
D	978	AAGAGGACCTCGGGGCTGGTGCCTGAGACCTGCAGAAAGAAAGAGGTGGAGTGGCTTTC	1033
Q	961	CAGTATTAATCAGTTACCAAGGTAAATATTTCCAATCACTCTGTGATGGAAGTGTGAC	1020
D	1038	CAGTATTAATCAGTTACCAAGGTAAATATTTCCAATCACTCTGTGATGGAAGTGTGAC	1097
Q	1021	GCAGGAACCTTCTTAACCACTGGGCTGCAATTCACCTGGCACTGTAAGAAAGCCAGGGCCATC	1086
D	1098	GCAGGAACCTTCTTAACCACTGGGCTGCAATTCACCTGGCACTGTAAGAAAGCCAGGGCCATC	1157
Q	1081	ATGAGAGAGGTTATGAGACCTGCTCAGGCCCTCAATATATCACTCAGAGCTGAGGCAATGGA	1146
D	1158	ATGAGAGAGGTTATGAGACCTGCTCAGGCCCTCAATATATCACTCAGAGCTGAGGCAATGGA	1217

Oy	1141	GAAGGACAGACATCAACTTTTGGATCCAAAGCTGGAGCTGGAGCCAGTCACTGGAT	1200
Db	1218	GAGGGACAGACATCAACTTTTGGATCCAAAGCTGGAGCTGGAGCCAGTCACTGGAT	1277
Oy	1201	GACTTATACAAAGTATTTCTTCTTCATCACTCCACGAGACACCATGACTGATGGAT	1266
Db	1278	GACTTATACAAAGTATTTCTTCTTCATCACTCCACGAGACACCATGACTGATGGAT	1337
Oy	1261	CCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTTGTATGTTGTGCACAC	1320
Db	1338	CCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTTGTATGTTGTGCACAC	1397
Oy	1321	ATGGAAGAAATGCTGCTAGTCC 1344	
Db	1398	ATGGAAGAAATGCTGCTAGTCC 1421	
RESULT 4			
AAV02296	AAV02296	standard; DNA; 1778 BP.	
AC	AAV02296;		
XX			
XX	21-MAY-1998	(first entry)	
XX			
XX	Human secreted protein AM282 full-length cDNA clone.		
XX			
XX	Secreted protein; AM282; cytokine; human; ds.		
XX			
XX	Homo sapiens.		
XX			
XX	Key	Location/Qualifiers	
XX	FT CDS	17..1435	
XX	FT	/*tag= a	
XX	FT	17..88	
XX	FT sig_peptide	/*tag= b	
XX	FT mat_peptide	89..1432	
XX	FT	/*tag= b	
XX			
XX	W09739030-A2.		
XX			
XX	23-OCT-1997.		
XX			
XX	16-APR-1997;	97WO-US06475.	
XX			
XX	13-JAN-1997;	97US-0783520.	
XX	PR 18-APR-1996;	96US-0634325.	
XX			
XX	(GEMV) GENETICS INST INC.		
XX			
XX	Jacobs K, LaVallie ER, McCoy JM, Merberg D, Racie LA;		
XX	PI Spaulding V;		
XX			
XX	WPI; 1997-526400/48.		
XX	DR P-PSDB; AAW33604.		
XX			
XX	New isolated secretory proteins AM340, AM282 and AK583 - possibly		
XX	PT have cytokine, cell proliferation/differentiation regulating,		
XX	PT Immunomodulating activities, etc.		
XX			
XX	Claim 15; Page 44-45; 59pp; English.		
XX			
XX	This cDNA clone encodes a protein (see W33604) designated AM282.		
XX	CC It was identified as "y95b10.r1 human EST 30142.5" (Genbank		
XX	CC accession No. R77830) in a database search using a partial AM282		
XX	CC clone (see T97398) obtained from a human foetal kidney cDNA		
XX	CC library using methods selective for cDNAs encoding secreted		
XX	CC proteins. AM282 is deposited in ATCC 98026 together with clones		
XX	CC AM340 (see T97397) and AK583 (see W02297), which are also claimed.		
XX	CC AM282 protein can be used in a claimed method for preventing,		
XX	CC treating or ameliorating a medical condition. It may exhibit		
XX	CC cytokine, cell proliferation (either inducing or inhibiting) or		

	Query Match	99.8%;	Score 1340.8;	DB 18;	Length 1778;	
	Best Local Similarity	99.9%;	Pred. No. 0;			
	Matches 1342;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
CC	cell differentiation (either inducing or inhibiting) activity or					
CC	may induce production of other cytokines in certain cell					
CC	populations. It may also exhibit e.g. immune stimulating or					
CC	suppressing activity, haematopoiesis regulating activity, tissue					
CC	growth activity, activin/inhibin activity, chemotactic or					
CC	chemokine/chemokine activity, haemostatic or thrombolytic activity.					
CC	receptor/ligand activity, anti-inflammatory activity, tumour					
CC	inhibition activity, or other activities. No evidence of any of					
CC	these activities is given in the specification.					
XX						
SQ	Sequence 1778 BP; 514 A; 386 C; 422 G; 456 T; 0 other;					
Query Match	99.8%;	Score 1340.8;	DB 18;	Length 1778;		
Best Local Similarity	99.9%;	Pred. No. 0;				
Matches 1342;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1 AAGATGGCATCTCTAAGAGAGACTTTTGAAGAAATTAAGAAATATGACAGCTGTGCA	60				
DB	89 AAGATGGCATCTCTAAGAGAGACTTTTGAAGAAATTAAGAAATATGACAGCTGTGCA	148				
QY	61 GATGTGGTAAAGCAATCATCAACCTGACTGTTTATGGTAAAGCCAGAACAGATCTTAT	120				
DB	149 GATGTGGTAAAGCAATCATCAACCTGACTGTTTATGGTAAAGCCAGAACAGATCTTAT	208				
QY	121 GAGGAGTTGGACATCTTGTTGTAATCTGTTGGAGCCAGACTGAGTGGTCCAGAACTTA	180				
DB	209 GAGGAGTTGGACATCTTGTTGTAATCTGTTGGAGCCAGACTGAGTGGTCCAGAACTTA	268				
QY	181 GAAAAAGCCATCCAAATTAATGATACCAAAACCTGACAGCAAGATGGGCTGGAAGAAAGTTAC	240				
DB	269 GAAAAAGCCATCCAAATTAATGATACCAAAACCTGACAGCAAGATGGGCTGGAAGAAAGTTAC	328				
QY	241 CTGGAGCCAGTGAAGATTAACCCACTGGAGAGAGGAGAGAAAGATCAGCTGTATGCTGGAG	300				
DB	329 CTGGAGCCAGTGAAGATTAACCCACTGGAGAGAGGAGAGAAAGATCAGCTGTATGCTGGAG	388				
QY	301 CCAAGAATTCTAAGATATGCCATCTTGAGTTTGAGCAGACAGATTGGGACTTCCTCAGAA	360				
DB	389 CCAAGAATTCTAAGATATGCCATCTTGAGTTTGAGCAGACAGATTGGGACTTCCTCAGAA	448				
QY	361 GGCAATTACAGAGAGATCTGGTGGGTACCCCTTTTCATGAACTGACAGAGAAAGGCGCTCA	420				
DB	449 GGCAATTACAGAGAGATCTGGTGGGTACCCCTTTTCATGAACTGACAGAGAAAGGCGCTCA	508				
QY	421 GAAACAGAGAGGAAGATGTTGTTTATTAACCAACTTAATCAATCACTCAAGACGCTG	480				
DB	509 GAAACAGAGAGGAAGATGTTGTTTATTAACCAACTTAATCAATCACTCAAGACGCTG	568				
QY	481 CAATACCGAACGACAGGGGGCGGTGAAAGCTGCCAAGGTGGGGGCTTTGGACTCTCTCAT	540				
DB	569 CAATACCGAACGACAGGGGGCGGTGAAAGCTGCCAAGGTGGGGGCTTTGGACTCTCTCAT	628				
QY	541 CGATCCGAGGCTCTCTCATCTTAACAGTCCCTCAACAGGTATTTCAAGAAATACAGGAT	600				
DB	629 CGATCCGAGGCTCTCTCATCTTAACAGTCCCTCAACAGGTATTTCAAGAAATACAGGAT	688				
QY	601 GGCGTGCACAAATATTCACACAGCCTGTATTAACGTGGAAGATGACGAAATGATGTCAAGA	660				
DB	689 GGCGTGCACAAATATTCACACAGCCTGTATTAACGTGGAAGATGACGAAATGATGTCAAGA	748				
QY	661 ATGCACTTCTCATGGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGGCAAGACCTTACCA	720				
DB	749 ATGCACTTCTCATGGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGGCAAGACCTTACCA	808				
QY	721 GATATCTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTT	780				
DB	809 GATATCTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTT	868				
QY	781 GTACTGTGACAGTGAACATCTGACAGAGCTGGGAGTGTGGGACAGGGTGCATGATGATGAC	840				
DB	869 GTACTGTGACAGTGAACATCTGACAGAGCTGGGAGTGTGGGACAGGGTGCATGATGATGAC	928				
QY	841 GGTGAGACTTTTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTTCA	900				

Db	929	GGTGGAGCCTTTATATCATGGAGACATCTTCATCTTATTTAAAGATCTTGGGCTGGCTCA	988
Qy	901	AAGAGACCTTCGCGGCTGCTGCTGGACTGCAAGAAACAAGTGGAGTTGGTCCCTTC	960
Db	989	AAGAGGACTCTGGCGGCTGGTGGCTCTGGACTGCAAGAAACAAGTGGAGTTGGTCCCTTC	1048
Qy	961	CAGTATTTATCAAGTTACACAAAGTAAATATTTCCACTACAGTCTGGTATGAGAGTCAAC	1022
Db	1049	CAGTATTTATCAAGTTACACAAAGTAAATATTTCCACTACAGTCTGGTATGAGAGTCAAC	1108
Qy	1021	GCAGGAACCTTCTTACCACTGGGCTGCACAATCATTGCAAGTGAAGAAAGCCAGGGCATC	1080
Db	1109	GCAGGAACCTTCTTACCACTGGGCTGCACAATCATTGCAAGTGAAGAAAGCCAGGGCATC	1160
Qy	1081	ATGAGAGAGTTATGAGCCTGCTGCAGCCCTCAATATATCACTCAGGTCTTGAGCCATGGA	1140
Db	1169	ATGAGAGAGTTATGAGCCTGCTGCAGCCCTCAATATATCACTCAGGTCTTGAGCCATGGA	1228
Qy	1141	GAAAGGACACAGACATCAACTTTTGGATCCAAAGTGGAGTGGCCCTGAGGCCAGTCACTACTGAT	1200
Db	1229	GAAAGGACACAGACATCAACTTTTGGATCCAAAGTGGAGTGGCCCTGAGGCCAGTCACTACTGAT	1288
Qy	1201	GACTATATCAAGTATTTCTTCTTCCATCACTCCCAAGGAGACACCATGACTGTCAATGAT	1260
Db	1289	GACTATATCAAGTATTTCTTCTTCCATCACTCCCAAGGAGACACCATGACTGTCAATGAT	1348
Qy	1261	CCAAAGCAGATGATATGTTGCTGCTGCTGTTGGAGCTGTTGTTCTTATATGTTGTCAGAC	1320
Db	1349	CCAAAGCAGATGATATGTTGCTGCTGCTGTTGGAGCTGTTGTTCTTATATGTTGTCAGAC	1408
Qy	1321	ATGGAAGAAATGCTCCCTTAGTCC	1344
Db	1409	ATGGAAGAAATGCTCCCTTAGTCC	1432

XX	AAZ98034	standard; CDNA; 1663 BP.
XX	AAZ98034;	
XX	09-MAY-2000	(first entry)
XX		
DE	Human secreted protein encoding nucleotide sequence SEQ ID NO:28.	
XX		
KW	Human; secreted protein; diagnosis; cytostatic; immunosuppressive;	
KW	antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;	
KW	tumour; neurodegenerative disorder; developmental abnormality; allergy;	
KW	foetal deficiency; blood disorder; immune system disorder; arthritis;	
KW	autoimmune disease; hepatic disease; renal disease; inflammation;	
KW	Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;	
KW	infection; AIDS; spinal cord injury; transplant rejection; diabetes;	
KW	asthma; sepsis; acne; psoriasis; cardiovascular disorder;	
KW	reproductive disorder; gastrointestinal disorder; respiratory disorder;	
KW	metabolic disorder; food additive; preservative; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200004140-A1.	
XX		
PD	27-JAN-2000.	
XX		
PF	14-JUL-1999;	99WO-US15849.
XX		
PR	15-JUL-1998;	98US-0092921.
PR	15-JUL-1998;	98US-0092922.
PR	15-JUL-1998;	98US-0092956.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y, Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;	

PI MucenSKI M., Endress GA, Soppet DR;
XX
XX WPI; 2000-161128/14.
DR
DR P-PSDB; MAY87081.
XX
XX New isolated human genes, useful for diagnosis and treatment of, e.g
PI
PI cancers, neurological or blood disorders -
XX
XX Claim 1; Page 319; 494pp; English.
PS

CC The polynucleotide sequences given in AA298017 to AA298108 encode the
CC human secreted proteins given in AA870064 to AA87223. Human secreted
CC protein can have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: cytostatic;
CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
CC antiallergic. The polynucleotides and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Human secreted protein s and their polynucleotides can
CC be used for developing products for the diagnosis or treatment of cancer,
CC tumours, neurodegenerative disorders, developmental abnormalities and
CC foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune disease, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The
CC proteins or polynucleotides can also be used as food additives or
CC preservatives. The proteins are also useful for identifying their
CC binding partners. AA288008 to AA288016 and AA870063 are sequence used in
CC the exemplification of the present invention.

sq' sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match	99.8%;	Score 1340.8;	DB 21;	Length 1863;
-------------	--------	---------------	--------	--------------

Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	AAGAATGGCATCTCTTAAGAGGACCTTTGAAGAAATTAAGAAAGAAATTAAGCAGCTGTGGA	60
Db	171	AAGAATGGCATCTCTTAAGAGGACCTTTGAAGAAATTAAGAAAGAAATTAAGCAGCTGTGGA	230
QY	61	GATGTGCTAAAGCAATCATCAACCTAGCTGTTATAGTAAAGCCAGAAACAGATCTTAT	120
Db	231	GATGTGCTAAAGCAATCATCAACCTAGCTGTTATAGTAAAGCCAGAAACAGATCTTAT	290
QY	121	GACGATTTGGCACTTCTGGTTGATATCTGTGGACCCAGATCGATAGCTGCTCAAGAACTTA	180
Db	291	GAGGAGTTGGCACTTCTGGTTGATATCTGTGGACCCAGATCGATAGCTGCTCAAGAACTTA	350
QY	181	GAAAAAGCCATCCCAATTATGTATACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTTAC	240
Db	351	GAAAAAGCCATCCCAATTATGTATACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTTAC	410
QY	241	CTGGAGCCAGTGAAGATTAACCCACTGGGAGAGAGGGAGAAAGATCACTGTGTGATGCTTGAG	300
Db	411	CTGGAGCCAGTGAAGATTAACCCACTGGGAGAGAGGGAGAAAGATCACTGTGTGATGCTTGAG	470
QY	301	CCAAGAATTATATAGATAGCCATCTCGGGTCTTTGCACAGACATTTGGGACTCTCCAGAA	360
Db	471	CCAAGAATTATATAGATAGCCATCTCGGGTCTTTGCACAGACATTTGGGACTCTCCAGAA	530
QY	361	GGCATTTACAGACAGAACTTCTGCTGTGAACCTTTTCATGAACTGCAGAGAAAGGGCTCA	420
Db	531	GGCATTTACAGACAGAACTTCTGCTGTGAACCTTTTCATGAACTGCAGAGAAAGGGCTCA	590
QY	421	GAAGCAAGAGGAAAGTTGTGTTTATATACCACTTATCATCAACTACTCAAGAGCGGTG	480
Db	591	GAAGCAAGAGGAAAGTTGTGTTTATATACCACTTATCATCAACTACTCAAGAGCGGTG	650

QY 481 CAATACGAGCGAGGGGGCGGTGGAGAGCTCCAAAGTGGGGCTTTGGCATCTCTCAT 540
 Db 651 CAATACGAGCGAGGGGGCGGTGGAGAGCTCCAAAGTGGGGCTTTGGCATCTCTCAT 710
 QY 541 CGATCCGCGGCT 600
 Db 711 CGATCCGCGGCT 770
 QY 601 GCGCGTCCCAAAATTTCCAAAGAGCTGTATTAAGGTGAAGATGCAAAATGATGCAAGA 660
 Db 771 GCGCGTCCCAAAATTTCCAAAGAGCTGTATTAAGGTGAAGATGCAAAATGATGCAAGA 830
 QY 661 ATGGCTTCTCATGGGATCAAAATTTGTCTATTCAGCTAAAGATGGGGGCAAGACCTACCA 720
 Db 831 ATGGCTTCTCATGGGATCAAAATTTGTCTATTCAGCTAAAGATGGGGGCAAGACCTACCA 890
 QY 721 GATCTGATTTCTCTCAACACTGTGTAGCAGAGATCACTGGGAGCAAAATATCCAAAGGTT 780
 Db 891 GATCTGATTTCTCTCAACACTGTGTAGCAGAGATCACTGGGAGCAAAATATCCAAAGGTT 950
 QY 781 GTACTGGTCAGTGGACATCTGGAAGAGCTGGGATGTTGGGGAGGCCATGATGATGAGC 840
 Db 951 GTACTGGTCAGTGGACATCTGGAAGAGCTGGGATGTTGGGGAGGCCATGATGATGAGC 1010
 QY 841 GGTGAGAGCTTTATATCATATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCA 900
 Db 1011 GGTGAGAGCTTTATATCATATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCA 1070
 QY 901 AAGAGAGCTCTGCGGCTGTGTCTCTGAGCTGAGAGAGAGCAAGGTGGTGGCTTC 960
 Db 1071 AAGAGAGCTCTGCGGCTGTGTCTCTGAGCTGAGAGAGAGCAAGGTGGTGGCTTC 1130
 QY 961 CAGTATTAATCAGTTACCAAGATTAATTTTCAACTTCACTGTGTGATGAGAGCTGAGC 1020
 Db 1131 CAGTATTAATCAGTTACCAAGATTAATTTTCAACTTCACTGTGTGATGAGAGCTGAGC 1190
 QY 1021 GCAGGAACCTTTCTTACCACCTGAGGCTGCATATTCATCTGAGTGAAGAGCCAGCCATC 1080
 Db 1191 GCAGGAACCTTTCTTACCACCTGAGGCTGCATATTCATCTGAGTGAAGAGCCAGCCATC 1250
 QY 1081 ATGAGAGAGTTTATGAGCTGTGTCTGAGCTGAGAGAGAGCAAGGTGGTGGCTTC 1140
 Db 1251 ATGAGAGAGTTTATGAGCTGTGTCTGAGCTGAGAGAGAGCAAGGTGGTGGCTTC 1310
 QY 1141 GAAGGAGCAGACATCACTTTGATCCAGCTGAGAGCTGAGAGAGCTGATCTTATGAT 1200
 Db 1311 GAAGGAGCAGACATCACTTTGATCCAGCTGAGAGCTGAGAGAGCTGATCTTATGAT 1370
 QY 1201 GACTTATTAAGATTTCTTCTCCATCACTCCACGAGAGACCATGATCTGATGAT 1260
 Db 1371 GACTTATTAAGATTTCTTCTCCATCACTCCACGAGAGACCATGATCTGATGAT 1430
 QY 1261 CCAAGAGAGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1330
 Db 1431 CCAAGAGAGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1490
 QY 1321 ATGGAAGAAATGCTGCTGCTGCTC 1344
 Db 1491 ATGGAAGAAATGCTGCTGCTGCTC 1514

RESULT 6

AAD11647 standard; cDNA; 1863 BP.

AAD11647;

24-SEP-2001 (first entry)

Human secreted protein-encoding gene 18 cDNA clone HRA0315, SEQ ID NO:28.
 Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 foetal abnormality; developmental abnormality; haematopoietic disorder;

KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;
 KW angiotensin disorder; kidney disorder; gastrointestinal disorder; allergy;
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KW cell culture; chemotaxis; vulvexery; binding partner identification;
 KW gene therapy; ss.
 KW OS
 KW Homo sapiens.
 XX
 FH Key
 FT CDS
 FT Location/Qualifiers
 FT 99..1517
 FT /*tag= a
 FT /product= "Human secreted protein precursor"
 FT sig_peptide
 FT 99..170
 FT /*tag= b
 FT mat_peptide
 FT 171..1514
 FT /*tag= c
 FT /product= "Mature human secreted protein"
 XX
 PN W0200151504-A1.
 XX
 PD 19-JUL-2001.
 XX
 PF 12-JAN-2001; 2001WO-US00911.
 XX
 PR 13-JAN-2000; 2000US-0482273.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Komatsu JI GA, Duan DR, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Endress GA, Muscenski M, Ebner R;
 XX
 DR WPI: 2001-425865/45.
 XX
 PT P-PSDB; AAE06058.
 XX
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 XX
 Claim 1; Page 686; 864pp; English.
 CC AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
 CC protein genes, and AAE06041-AAE06132 represent the proteins they encode.
 CC AAE06133-AAE06205 represent human secreted protein fragments.
 CC The secreted proteins and their genes are useful for preventing, treating
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 CC Pathological conditions can be diagnosed by determining the amount of the
 CC new protein in a sample or by determining the presence of mutations in
 CC the new genes. Specific uses are described for each of the 71 genes,
 CC based on the tissues in which they are most highly expressed, and include
 CC developing products for the diagnosis or treatment of proliferative
 CC disorders, cancer, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
 CC pregnancy-related disorders, endocrine disorders, and infections. The
 CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, for supporting cell culture of primary tissues,
 CC to regenerate tissues, to identify their cognate ligands or binding
 CC partners, and in chemotaxis, and can be used as a food additive or
 CC preservative to modify storage properties. Antibodies specific for a
 CC protein of the invention can be used in alleviating symptoms associated
 CC with the disorders mentioned above, and in diagnostic immunoassays (e.g.,
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA)).
 CC The present sequence represents a human secreted protein-encoding cDNA of
 CC the invention.
 XX

SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 99.8%; Score 1340.8; DB 22; Length 1863;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AAGATGCGATCTCTAAGAGGACCTTTGAGAAATTAAGAAATAGCCAGCTGTGGA 60
171 AAGATGCGATCTCTAAGAGGACCTTTGAGAAATTAAGAAATAGCCAGCTGTGGA 230
61 GATGTTGCTAAAGCAATCATCACTAGCTGTTATATGTTAAAGCCAGAACGATCTAT 120
221 GATGTTGCTAAAGCAATCATCACTAGCTGTTATATGTTAAAGCCAGAACGATCTAT 290
121 GAGCATTTGGCACTCTGTTGATATCTGTGGAACCAAGTGAAGTGGTCTCCAGAACTTA 180
291 GAGCATTTGGCACTCTGTTGATATCTGTGGAACCAAGTGAAGTGGTCTCCAGAACTTA 350
181 GAAAAAGCCATCCAAATTAATGTAACAAAACTGACAGAAATGGGCTGGAGAAATTTAC 240
351 GAAAAAGCCATCCAAATTAATGTAACAAAACTGACAGAAATGGGCTGGAGAAATTTAC 410
241 CTGAGCAGATGAGAAATACCCCACTGGAGAGAGGAGAAATTCAGCTGTGAG 300
411 CTGAGCAGATGAGAAATACCCCACTGGAGAGAGGAGAAATTCAGCTGTGAG 470
301 CCAAGAAATTCATTAATAGATCCATCTGAGTCTTGGCAGCAGATTTGGATCTCTCAGAA 360
471 CCAAGAAATTCATTAATAGATCCATCTGAGTCTTGGCAGCAGATTTGGATCTCTCAGAA 530
361 GGCATTACAGCAGAAATCTGAGTGTGACCTCTTTCAGATGAACTGACAGAAAGGCTCA 420
531 GGCATTACAGCAGAAATCTGAGTGTGACCTCTTTCAGATGAACTGACAGAAAGGCTCA 590
421 GAAGCAAGAGGAGATGTTGTTTATTAACAACCTTACATCACTACTCAAGACGGTG 480
591 GAAGCAAGAGGAGATGTTGTTTATTAACAACCTTACATCACTACTCAAGACGGTG 650
481 CAATCCCAAGCAGAGGAGGCGGTGAGAGCTGCAAGTGGGGCTTTGGCATCTCTCAT 540
651 CAATCCCAAGCAGAGGAGGCGGTGAGAGCTGCAAGTGGGGCTTTGGCATCTCTCAT 710
541 CGATCCGAGGCTCTCTCCATCTACAGTCTCTCACAAGATTCAGAAATACAGAGAT 600
711 CGATCCGAGGCTCTCTCCATCTACAGTCTCTCACAAGATTCAGAAATACAGAGAT 770
601 GGCCTGCCCAAAATTCACAAGCTGTATTAACGTTGAGAAATGACAAATATGTCAGA 660
771 GGCCTGCCCAAAATTCACAAGCTGTATTAACGTTGAGAAATGACAAATATGTCAGA 830
661 ATGCTTCTCATGGGATCAAAATTTGTCTCATTCAGCTTAAAGATGGGGCAAAAGCTTACCA 720
831 ATGCTTCTCATGGGATCAAAATTTGTCTCATTCAGCTTAAAGATGGGGCAAAAGCTTACCA 890
721 GATACGATCTCTTCAACAGCTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAAGGTT 780
891 GATACGATCTCTTCAACAGCTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAAGGTT 950
781 GTACTGTCAGTGAATCTGAGCAGCTGGAGATGTTGGGACAGGTCAGATGATGGC 840
951 GTACTGTCAGTGAATCTGAGCAGCTGGAGATGTTGGGACAGGTCAGATGATGGC 1010
841 GGTGAGGCTTTATATCATGGGAAGCAGCTCTCATTTTAAAGATCTTGGGCTGGCTCCA 900
1011 GGTGAGGCTTTATATCATGGGAAGCAGCTCTCATTTTAAAGATCTTGGGCTGGCTCCA 1070
901 AAGAGAGCTCTGCGCTGCTGAGCTGACAGAGAAACAAGTGGATGCTGCTTC 960
1071 AAGAGAGCTCTGCGCTGCTGAGCTGACAGAGAAACAAGTGGATGCTGCTTC 1130
961 CAGTATTAACAGTTACACAAGGTAATATTTTCAACTACAGTCTGATGAGAGTCTGAC 1020
1131 CAGTATTAACAGTTACACAAGGTAATATTTTCAACTACAGTCTGATGAGAGTCTGAC 1190

1021 GCAGAACTTCTTACCACTGGGCTGCAATTCAGTGGAGTAAAGCCAGGCGCATC 1080
1191 GCAGAACTTCTTACCACTGGGCTGCAATTCAGTGGAGTAAAGCCAGGCGCATC 1250
1081 ATGAGAGAGGTTATGAGCTGCTGACAGCCCTCAATATCACTCAGTCTGAGCCATGA 1140
1251 ATGAGAGAGGTTATGAGCTGCTGACAGCCCTCAATATCACTCAGTCTGAGCCATGA 1310
1141 GAAGGACAGACATCACTTTTGGATCCAGCTGAGTGGCTGAGCCAGTCTTATGAT 1200
1311 GAAGGACAGACATCACTTTTGGATCCAGCTGAGTGGCTGAGCCAGTCTTATGAT 1370
1201 GACTTATACAGATTTTCTTCTTCATCACTCCACAGGAGACACATGATCATGAT 1260
1371 GACTTATACAGATTTTCTTCTTCATCACTCCACAGGAGACACATGATCATGAT 1430
1261 CCAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTTGTTCTTATGTTGTCAGAC 1320
1431 CCAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTTGTTCTTATGTTGTCAGAC 1490
1321 ATGGAAGAAATGCTGCTGCTGCTC 1344
1491 ATGGAAGAAATGCTGCTGCTGCTC 1514

RESULT 7
ABK69743
ID ABK69743 standard; cDNA; 1863 BP.
XX AC ABK69743;
XX DT 15-JUL-2002 (first entry)
XX DE Human secreted protein gene 18 #1.
XX KW Human; sex; gene; secreted protein; gene therapy; immunosuppressive;
KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
KW vasotrophic; cerebroprotective; nootropic; neuroprotective; antibacterial;
KW vitruclide; fungicide; ophthalmological; autoimmune disease; neoplasm;
KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing; skin aging;
KW epithelial cell proliferation; food additive.
XX OS Homo sapiens.
XX PN WO200226931-A2.
XX PD 04-APR-2002.
XX PF 24-SEP-2001; 2001WO-US29871.
XX PR 25-SEP-2000; 2000US-234925P.
XX PR 12-JAN-2001; 2001WO-US00911.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y,
PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
PI Endreess GA, Mucenski M, Ebner R;
XX DR WPI; 2002-362489/39.
XX DR P-PSDB; ABG33880.
XX PT Novel 71 isolated secreted polypeptides and polynucleotides encoding
PT the polypeptides, useful for treating Huntington's disease, sepsis,
PT meningitis, thrombocytopaenia, haemolytic anaemia, rheumatoid arthritis,
PT asthma
XX PS Claim 1; Page 1170; 1478pp; English.
XX

CC The invention relates to an isolated nucleic acid molecule (or its
 CC fragment, homologue fragment or allele variant) encoding a human
 CC secreted protein (and its fragment, domain, epitope, variant, secreted
 CC form and species variant). Also included are a recombinant vector
 CC comprising the nucleic acid, a recombinant host cell comprising the
 CC vector, an antibody against the secreted protein, a recombinant host cell
 CC that expresses the secreted protein and a method of identifying a binding
 CC partner of the secreted protein. The nucleic acid and protein are used to
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
 CC human, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
 CC for example autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiodenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
 CC ocular disorders e.g. corneal infection. Many other diseases and
 CC disorders are listed in the specification. The polypeptides can also be
 CC used to aid wound healing an epithelial cell proliferation, to prevent
 CC skin aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities. The present
 CC sequence encodes a novel human secreted protein of the invention.

XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 99.8%; Score 1340.8; DB 24; Length 1863;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 134; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGATGGCATCTCTAAGAGACCTTTGAAGAAATTAAGAAATAGCCAGCTGTGA 60
 DB 171 AAGATGGCATCTCTAAGAGACCTTTGAAGAAATTAAGAAATAGCCAGCTGTGA 230
 QY 61 GATGTTCTAAGCAATCATCACTAGCTGTTATAGTAAAGCCAGACATCTAT 120
 DB 231 GATGTTCTAAGCAATCATCACTAGCTGTTATAGTAAAGCCAGACATCTAT 290
 QY 121 GAGGATTTGGCATCTGTTGATGATGTTGACCCAGACTGAGTGGTCCCAAGACTA 180
 DB 291 GAGGATTTGGCATCTGTTGATGATGTTGACCCAGACTGAGTGGTCCCAAGACTA 350
 QY 181 GAAAAAGCATCCAAATTAATGATCAAAACCTGAGCAAGATGGGCTGAGAAAGTTAC 240
 DB 351 GAAAAAGCATCCAAATTAATGATCAAAACCTGAGCAAGATGGGCTGAGAAAGTTAC 410
 QY 241 CTGGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAAGATCAGCTGTGATGCTGAG 300
 DB 411 CTGGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAAGATCAGCTGTGATGCTGAG 470
 QY 301 CCAAGATTTATAGATAGCATCTGAGTCTTGGAGCAGCATTTGGGACTCTCCAGAA 360
 DB 471 CCAAGATTTATAGATAGCATCTGAGTCTTGGAGCAGCATTTGGGACTCTCCAGAA 530
 QY 361 GGCATTACAGAGAGTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 420
 DB 531 GGCATTACAGAGAGTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 590
 QY 421 GAAGCAAGAGAGAGTTGTTGTTATTAACCACTTAATCACTCAAGAGAGGCTG 480
 DB 591 GAAGCAAGAGAGAGTTGTTGTTATTAACCACTTAATCACTCAAGAGAGGCTG 650
 QY 481 CAATACCGAAGCGAGGGGGGGGGGAGAGCTGCAAGGTGGGGCTTTGGCATCTCAT 540
 DB 651 CAATACCGAAGCGAGGGGGGGGGGAGAGCTGCAAGGTGGGGCTTTGGCATCTCAT 710
 QY 541 CGATCCGTTGGCTCTTCTCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAT 600
 DB 711 CGATCCGTTGGCTCTTCTCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAT 770
 QY 601 GGCCTGCCCAAAATTTCCAAAGCTGTATTAAGGTGAAGATGCAAAATGATGTCAAGA 660
 DB 771 GGCCTGCCCAAAATTTCCAAAGCTGTATTAAGGTGAAGATGCAAAATGATGTCAAGA 830

QY 661 ATGGCTTCTCATGGGATCAAAATTTGTCATTAGCTAAAGATGGGGGCAAGACTTACCA 720
 DB 831 ATGGCTTCTCATGGGATCAAAATTTGTCATTAGCTAAAGATGGGGGCAAGACTTACCA 890
 QY 721 GATATGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGTT 780
 DB 891 GATATGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGTT 950
 QY 781 GATCTGCTAGTGAACATCTGACAGCTGGAGTCTTGGGAGGCTGCAAGATGATGCG 840
 DB 951 GATCTGCTAGTGAACATCTGACAGCTGGAGTCTTGGGAGGCTGCAAGATGATGCG 1010
 QY 841 GGTGAGACCTTTTATATCATGGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCTCA 900
 DB 1011 GGTGAGACCTTTTATATCATGGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCTCA 1070
 QY 901 AAGAGACTCTGCGGCTGTGCTCTGAGCTGCAAGAGAACAGGTGAGTTGCTCTTC 960
 DB 1071 AAGAGACTCTGCGGCTGTGCTCTGAGCTGCAAGAGAACAGGTGAGTTGCTCTTC 1130
 QY 961 CAGTATTTATAGTTACCAAGGTTAAATTTTCCATCAAGTCTGTTGATGAGTCTGAC 1020
 DB 1131 CAGTATTTATAGTTACCAAGGTTAAATTTTCCATCAAGTCTGTTGATGAGTCTGAC 1190
 QY 1021 GCAAGAACTTTTACCCTGAGCTGCAATTCAGTGGAGTGAAGGAGGAGGCTATC 1080
 DB 1191 GCAAGAACTTTTACCCTGAGCTGCAATTCAGTGGAGTGAAGGAGGAGGCTATC 1250
 QY 1081 ATGAGAGAGTTATGAGCTGCTGACAGCCCTCAATATCACTAGTCTGAGGCAATGA 1140
 DB 1251 ATGAGAGAGTTATGAGCTGCTGACAGCCCTCAATATCACTAGTCTGAGGCAATGA 1210
 QY 1141 GAGGAGCAAGCATCACTTTTGGATCAAGCTGAGTGGCTGAGGCAATCTTATGAT 1200
 DB 1311 GAGGAGCAAGCATCACTTTTGGATCAAGCTGAGTGGCTGAGGCAATCTTATGAT 1260
 QY 1201 GACTTATCAAGATTTCTTCTTCCATCACTCCAGAGAGACCAATGATGATGAT 1260
 DB 1371 GACTTATCAAGATTTCTTCTTCCATCACTCCAGAGAGACCAATGATGATGAT 1320
 QY 1261 CCAAGCAGATGATGTTGCTGCTGCTGTTGAGGCTGTTTCTTATGTTGTCAGAC 1320
 DB 1431 CCAAGCAGATGATGTTGCTGCTGCTGTTGAGGCTGTTTCTTATGTTGTCAGAC 1380
 QY 1321 ATGGAAGAAATGCTGCTAGGCTC 1344
 DB 1491 ATGGAAGAAATGCTGCTAGGCTC 1514

RESULT 8
 ACCS0817
 ID ACCS0817 standard; cDNA; 1863 BP.
 XX ACCS0817;
 AC ACCS0817;
 DT 12-JUN-2003 (first entry)
 XX
 DE Human secreted protein coding sequence, SEQ ID 484.
 XX
 KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;
 KW valnetary; antiinflammatory; nootropic; neuroprotective;
 KW antiParkinsonian; gene therapy; human; cardiovascular disorder;
 KW gene; ss.
 XX
 OS Homo sapiens.
 PN W0200295010-A2.
 XX W0200295010-A2.
 PD 28-NOV-2002.
 XX 28-NOV-2002.
 PF 19-MAR-2002; 2002MO-US09785.
 XX 19-MAR-2002; 2002MO-US09785.

DT 04-APR-2003 (first entry)
XX Secreted protein-encoding gene 142 cDNA clone HRACJ35, SEQ ID NO:274.
XX
XX Human; secreted protein; digestive disorder; gastrointestinal disorder;
XX mouth; oesophagus; stomach; small intestine; large intestine; liver;
XX biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
XX immune disorder; inflammation; infection; wound healing; drug screening;
XX chromosome identification; chromosome mapping; cytostatic; gene therapy;
XX antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2;
XX gene; 88.
XX Homo sapiens.
XX WO200276488-A1.
XX
XX 03-OCT-2002.
XX
XX 19-MAR-2002; 2002WO-US08276.
XX
XX 21-MAR-2001; 2001US-277340P.
XX 19-JUL-2001; 2001US-306171P.
XX 13-NOV-2001; 2001US-331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-029900/02.
XX P-PSDB; ABR00274.
XX
XX New human secreted proteins and nucleic acids, useful for detecting,
XX preventing, diagnosing, prognosticating, treating and/or ameliorating
XX e.g. gastrointestinal diseases and disorders, or cancers -
XX
XX Claim 21; Page 918; 1216pp; English.
XX
XX AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
XX protein genes, and ABP00011-ABP00299 represent the proteins they encode.
XX AB271479-AB271540 represent human secreted protein genomic fragments. The
XX invention also encompasses antibodies specific for the secreted proteins,
XX the use of the secreted proteins in drug screening, and recombinant
XX vectors and host cells comprising a nucleic acid of the invention. The
XX secreted proteins, nucleic acids encoding them, antibodies or antibody
XX fragments specific for the secreted proteins, and modulators of protein
XX activity are useful for diagnosing, treating, ameliorating or preventing
XX digestive disorders. Such conditions include disorders of the mouth,
XX oesophagus, stomach, small intestine, large intestine, liver, biliary
XX tract and pancreas, and include cancers of these organs and tissues. The
XX secreted proteins and their nucleic acids may also be used in the
XX treatment of immune disorders, inflammation, infection,
XX hyperproliferative disorders, and to promote wound healing. Nucleic acids
XX of the invention may be used for chromosome identification, chromosome
XX mapping, in gene therapy, for identifying individuals from minute
XX biological samples, as hybridisation probes, and as molecular weight
XX markers. The present sequence represents a human secreted protein-
XX encoding cDNA clone of the invention.
XX
XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
SQ
Query Match 99.8%; Score 1340.8; DB 25; Length 1863;
Best Local Similarity 99.9%; Pred. No. 0; Indels 0; Gaps 0;
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAGATGGCATCTTAAGAGGACTTTTGAAGAAATTAAGAAATGAGCAGCTGTGA 60
DB 171 AAGATGGCATCTTAAGAGGACTTTTGAAGAAATTAAGAAATGAGCAGCTGTGA 230
QY 61 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCCAAGATCTTAT 120
DB 231 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCCAAGATCTTAT 290
QY 121 GAGGATTGGCACTTCTGTTGATGATGTTGGAACCAAGATGAGTGTCCAGAACTTA 180

DB 291 GAGGATTGGCACTTCTGTTGATGATGTTGGAACCAAGATGAGTGTCCAGAACTTA 350
QY 181 GAAAGACCATCCAAATTAATGATCAAAACCTGAGCAAGATGGGCTGGAGAAAGTTTAC 240
DB 351 GAAAGACCATCCAAATTAATGATCAAAACCTGAGCAAGATGGGCTGGAGAAAGTTTAC 410
QY 241 CTGAGGACAGTGAAGATATCCCACTGGAGAGGGGAGAGAAATAGCTGTGATCTGAG 300
DB 411 CTGAGGACAGTGAAGATATCCCACTGGAGAGGGGAGAGAAATAGCTGTGATCTGAG 470
QY 301 CCAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 471 CCAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 530
QY 361 GGCATTACAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 531 GGCATTACAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 590
QY 421 GAAGCAAGAGGAGAAATGTTGTTTATTAACCAACTTACATCACTCAAGAGAGGCTG 480
DB 591 GAAGCAAGAGGAGAAATGTTGTTTATTAACCAACTTACATCACTCAAGAGAGGCTG 650
QY 481 CAATACCGAAACGAGGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 651 CAATACCGAAACGAGGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710
QY 541 CGATCCGTCGCT 600
DB 711 CGATCCGTCGCT 770
QY 601 GGCCTGCGCCAAATTCGAACAGCTGATTAACGCTGAGAGAGAGAGAGAGAGAGAGAG 660
DB 771 GGCCTGCGCCAAATTCGAACAGCTGATTAACGCTGAGAGAGAGAGAGAGAGAGAGAG 830
QY 661 ATGCTTCTCATGGATCAAAATTTGTCATTCAGCTAAAGATGGGGGCAAGACCTACCA 720
DB 831 ATGCTTCTCATGGATCAAAATTTGTCATTCAGCTAAAGATGGGGGCAAGACCTACCA 890
QY 721 GATATGATATTCCTTAACACTGTAAGAGATCACTGGAGCAAAATTAACAGAGATT 780
DB 891 GATATGATATTCCTTAACACTGTAAGAGATCACTGGAGCAAAATTAACAGAGATT 950
QY 781 GATATGATATTCCTTAACACTGTAAGAGATCACTGGAGCAAAATTAACAGAGATT 840
DB 951 GATATGATATTCCTTAACACTGTAAGAGATCACTGGAGCAAAATTAACAGAGATT 1010
QY 841 GGTGAGACCTTTATATATATGAGGAAAGCACTCTCACTTAATTAAGATCTTGGCTGCTCA 900
DB 1011 GGTGAGACCTTTATATATATGAGGAAAGCACTCTCACTTAATTAAGATCTTGGCTGCTCA 1070
QY 901 AAGAGACTCTGCGGCTGTGCTGTGAGCTGACAGAGAAACAGAGAGAGTGTGCTTC 960
DB 1071 AAGAGACTCTGCGGCTGTGCTGTGAGCTGACAGAGAAACAGAGAGAGTGTGCTTC 1130
QY 961 CAGATTAATGATTAACCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
DB 1131 CAGATTAATGATTAACCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1190
QY 1021 GCAGGAACCTTTCTTAACCACTGGGCTGCAATTCAGCTGAGTGAAGGAGGAGGAGGATC 1080
DB 1191 GCAGGAACCTTTCTTAACCACTGGGCTGCAATTCAGCTGAGTGAAGGAGGAGGAGGATC 1250
QY 1081 ATGAGAGAGGTTATGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1251 ATGAGAGAGGTTATGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1310
QY 1141 GAAAGGACAGACATCACTTTTGAATCCAGAGCTGAGTGTGAGAGAGAGAGAGAGAGAG 1200
DB 1311 GAAAGGACAGACATCACTTTTGAATCCAGAGCTGAGTGTGAGAGAGAGAGAGAGAGAG 1370
QY 1201 GACTTATACAAATGATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260

Db	1371	GACCTATFACAAAGTATTCTTCTTCATGCATCCACGGAGACACATCATGTCATGGAT	14320
Qy	1261	CCAAAGCAGATGATGATGGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGGACAGC	13220
Db	1431	CCAAAGCAGATGAAATGCTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGGACAGC	14930
Qy	1321	ATGGAAGAAATGCTGCTTAGGCC	1344
Db	1491	ATGGAAGAAATGCTGCTTAGGCC	1514
RESULT 10			
ID	AAZ58313	standard; cDNA; 1884 BP.	
XX	AAZ58313;		
AC	AAZ58313;		
DT	08-MAY-2000	(first entry)	
XX			
DE	Human peptidase NALAD-ase IV cDNA.		
XX	NALAD-ase IV, N-acetylated alpha-linked acidic dipeptidase; human;		
KM	chromosome 8q21.3; prostate cancer; neurodegenerative disease;		
KM	Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;		
KM	peripheral neuropathy; Huntington's disease; acute brain injury;		
KM	multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;		
KM	gene therapy; diagnosis; nootropic; neuroprotective; neuroleptic;		
KM	antiparkinsonian; anticonvulsant; vasotropic; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	149..1567	
FT		/tag= a	
XX			
PN	WO200004157-A2.		
PD	27-JAN-2000.		
XX			
PF	14-JUL-1999;	99WO-GB02241.	
XX			
PR	14-JUL-1998;	98GB-0015284.	
XX			
PA	(JANC) JANSEN PHARM NV.		
PI	Pangalos M, Neefs JEFM, Peeters DCG;		
XX			
DR	WPI; 2000-182424/16.		
DR	P-PSDB; AAZ58879.		
XX			
PT	New human N-acetylated alpha-linked acidic dipeptidases for treating		
PT	neural disorders e.g. Alzheimer's disease, schizophrenia and		
PT	Parkinson's disease		
XX			
PS	Claim 7, Fig 5, 95pp; English.		
XX			
CC	The present sequence is that of cDNA coding for human		
CC	N-acetylated alpha-linked acidic dipeptidase IV (NALAD-ase IV,		
CC	see AAZ58879). The cDNA was obtained from a gall bladder cDNA		
CC	library. Analysis of the open reading frame predicts a type II		
CC	integral membrane protein with 5 potential N-glycosylation sites.		
CC	The NALAD-ase II gene was mapped to chromosome 8q21.3.		
CC	NALAD-ase IV expression was low in all tissues examined by RT-PCR.		
CC	The invention provides human NALAD-ase I, II and IV polypeptides,		
CC	CDNAs, antisense nucleic acids, vectors, host cells, transgenic		
CC	organisms, antagonists and agonists. These are useful for treating		
CC	neural disorders such as Alzheimer's disease, schizophrenia, ALS,		
CC	Parkinson's disease, peripheral neuropathy, Huntington's disease,		
CC	acute brain injury, multiple sclerosis, exposure to neurotoxins,		
CC	peripheral nerve trauma, ischaemia or dementia (claimed). Nucleic		
CC	acids can also be used for gene therapy and for genetic screening		
CC	of predisposition to disorders associated with NALAD-ase.		
XX			

Seq	Sequence	884 BP; 520 A; 432 C; 458 G; 474 T; 0 other;
QY	Query Match	99.8%; Score 1340.8; DB 21; Length 1884;
Db	Best Local Similarity	99.9%; Pred. No. 0;
QY	Matches 1342; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1 AAGAATGGCACTCTCTAAGAGGACTTTTGAAGAAATTAAGAAGAAATAGCCAGCTGGGA	60
Db	221 AAGAATGGCACTCTCTAAGAGGACTTTTGAAGAAATTAAGAAGAAATAGCCAGCTGGGA	280
QY	61 GATGTGGTAAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGACAGATCTAT	120
Db	281 GATGTGGTAAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGACAGATCTAT	340
QY	121 GAGGAGTTGGCACTTCTGGTTGATTAAGTCTGTTGAGACCTGACCTGAGTGGCTCCAAAGACTA	180
Db	341 GAGGAGTTGGCACTTCTGGTTGATTAAGTCTGTTGAGACCTGAGTGGCTCCAAAGACTA	400
QY	181 GAAAAAGCATCCAAATTTATGTACCAAAACCTGACGAAAGATGGGCTGGAGAAAGTTTAC	240
Db	401 GAAAAAGCATCCAAATTTATGTACCAAAACCTGACGAAAGATGGGCTGGAGAAAGTTTAC	460
QY	241 CTGGAGCCAGTGAAGAAATACCCCACTGGAGAGAGGAGAAAGATAGCTGTATCTGGAG	300
Db	461 CTGGAGCCAGTGAAGAAATACCCCACTGGAGAGAGGAGAAAGATAGCTGTATCTGGAG	520
QY	301 CCAAGAAATCTAAGATAGCCATCTGGGTCTTGGAGACAGCATTTGGGACTCTCTCGAA	360
Db	521 CCAAGAAATCTAAGATAGCCATCTGGGTCTTGGAGACAGCATTTGGGACTCTCTCGAA	580
QY	361 GGCATTACAGAGAAAGTTCTGGTGGTGAACCTCTTTGATGAACCTGACAGAAAGGGCCCTCA	420
Db	581 GGCATTACAGAGAAAGTTCTGGTGGTGAACCTCTTTGATGAACCTGACAGAAAGGGCCCTCA	640
QY	421 GAAGCAAGAGGGAAGATTTGTTTATTAACCACTTAACATCACTACTCAAGACGGT	480
Db	641 GAAGCAAGAGGGAAGATTTGTTTATTAACCACTTAACATCACTACTCAAGACGGT	700
QY	481 CAATACCGAAGCAGAGGGGCGGTGGAAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCAT	540
Db	701 CAATACCGAAGCAGAGGGGCGGTGGAAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCAT	760
QY	541 CGATTCGGTGGCTCTCTCTCATCTACAGTCTCTCAACAGGATTTACAGGAATACCAAGAT	600
Db	761 CGATTCGGTGGCTCTCTCTCATCTACAGTCTCTCAACAGGATTTACAGGAATACCAAGAT	820
QY	601 GGGCGGCCCAAAATTTCCAAAGCCTGTATTAAGGTGAAGATGACGAATGATGTCAAGA	660
Db	821 GGGCGGCCCAAAATTTCCAAAGCCTGTATTAAGGTGAAGATGACGAATGATGTCAAGA	880
QY	661 ATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGCTTAAAGATGGGGGCAAGACCTACCA	720
Db	881 ATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGCTTAAAGATGGGGGCAAGACCTACCA	940
QY	721 GATACTGATTCCTTCAACACTGTGACAGAGATCACTGGAGCAAAATATCCAAACAGGTT	780
Db	941 GATACTGATTCCTTCAACACTGTGACAGAGATCACTGGAGCAAAATATCCAAACAGGTT	1000
QY	781 GTACTGTGCACTGGACATCTGGACAGAGCTGGGATGTTGGGCGAGGGTGCATGGAATGATGGC	840
Db	1001 GTACTGTGCACTGGACATCTGGACAGAGCTGGGATGTTGGGCGAGGGTGCATGGAATGATGGC	1060
QY	841 GGTGGAGGCTTTATATCAATGGAAGCACTCACTTATTTAAAGATCTTTGGCTCGTCCA	900
Db	1061 GGTGGAGGCTTTATATCAATGGAAGCACTCACTTATTTAAAGATCTTTGGCTCGTCCA	1120
QY	901 AAGAGGACTTCGGCGCTGGTCTCTGGAATGACAGAAAGCAAGGTGAGATTGGTCTTTC	960
Db	1121 AAGAGGACTTCGGCGCTGGTCTCTGGAATGACAGAAAGCAAGGTGAGATTGGTCTTTC	1180
QY	961 CAGATTTATCATGTTATCACAGAGTAAATATTTTCCAACTACAGTCTGGGTGATGGAATCTGAC	1020
Db	1181 CAGATTTATCATGTTATCACAGAGTAAATATTTTCCAACTACAGTCTGGGTGATGGAATCTGAC	1240

QY	1021	GCAGGAACCTCTTCAACCACTGAGGCTGCATTTCACTGGCAGGAAAGGCCAGGGCCATC	1088
Db	1241	GCAGGAACCTCTTCAACCACTGAGGCTGCATTTCACTGGCAGGAAAGGCCAGGGCCATC	13000
QY	1081	ATGGAGGAGGTTATGAGGCTGCTGCAGGCCCTCAATATCACTCAAGTCTGAGCCATGGA	11440
Db	1301	ATGGAGGAGGTTATGAGGCTGCTGCAGGCCCTCAATATCACTCAAGTCTGAGCCATGGA	13660
QY	1141	GAAAGGACAGCATTCAACTTTTGGATCCAACTGAGAGTGGCTGGAGCCAGTCTACTTGAT	12000
Db	1361	GAAAGGACAGCATTCAACTTTTGGATCCAACTGAGAGTGGCTGGAGCCAGTCTACTTGAT	14200
QY	1201	GACTTATACAGATATTTCTTCTTCATCACTCACTCCACGAGGACACCATGACTGTATGAT	12660
Db	1421	GACTTATACAGATATTTCTTCTTCATCACTCACTCCACGAGGACACCATGACTGTATGAT	14860
QY	1261	CCAAAGCAGATGATATGTTGCTGCTGCTGTTTGGGCTGTTGTTTCTTATGTTGTCAGAC	13200
Db	1481	CCAAAGCAGATGATATGTTGCTGCTGCTGTTTGGGCTGTTGTTTCTTATGTTGTCAGAC	15400
QY	1321	ATGGAGAAGATGCTGCTTGAAGTCC	1344
Db	1541	ATGGAGAAGATGCTGCTTGAAGTCC	1564
RESULT 11			
AA298139			
ID	AA298139	standard; cDNA; 1923 BP.	
XX	AA298139;		
AC			
XX			
DT	11-MAY-2000	(first entry)	
DE	Human signal peptide containing protein HSP-31 cDNA SEQ ID NO:165.		
XX			
KW	Human; signal peptide-containing protein; HSP; diagnosis; cancer;		
KW	inflammation; cardiovascular disease; anticancer; anti-inflammation;		
KW	antimicrobial; neuroprotective; cardiovascular; hepatotropic;		
KW	antiaesthetic; gene therapy; cell proliferation; neurological disorder;		
KW	reproductive disorder; developmental disorder; arteriosclerosis;		
KW	clitrositis; psoriasis; acquired immune deficiency syndrome; anaemia;		
KW	asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;		
KW	Parkinson's disease; Huntington's diseases; ovulatory defect;		
KW	muscular dystrophy; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200000610-A2.		
XX			
PD	06-JAN-2000.		
XX			
PF	25-JUN-1999;	99WO-US14484.	
XX			
PR	26-JUN-1998;	98US-0090762.	
XX			
PR	31-JUL-1998;	98US-0094983.	
XX			
PR	01-OCT-1998;	98US-0102686.	
XX			
PR	11-DEC-1998;	98US-0112129.	
XX			
PA	(INCY-) INCYTE PHARM INC.		
XX			
PI	Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;		
XX			
PI	Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;		
XX			
PI	Bandman O;		
XX			
DR	WPI, 2000-160673/14.		
XX			
DR	P-PSDB; AAY87254.		
XX			
PT	New human signal peptide-containing proteins useful in treatment,		
XX			
PT	prevention and diagnosis of e.g. cancer, inflammation and		
XX			
PT	cardiovascular disease		
XX			
XX	Claim 9, Page 269-270, 327pp, English.		

Query Match	Best Local Similarity	99.8%	Score 1340.8	DB 21	Length 1923
Matches 1342	Conservative	0	Mismatches 2	Indels 0	Gaps 0
XX	AA298109	to AA298242	encode AAY87224	to AAY87357	which represent the human signal peptide-containing proteins HSPB-1 to HSPB-134. HSPBs have anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic, neuroprotective, cardiovascular and antistatic activities, and can be used in gene therapy. HSPBs can be used to treat or prevent disorders associated with decreased activity or function of HSPB. Antagonists of HSPB are used to treat or prevent disorders associated with increased activity or function of HSPB. Such disorders include cell proliferation (including cancer), inflammation, cardiovascular, neurological, reproductive or developmental disorders (e.g. arteriosclerosis, cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia, asthma, Crohn's disease, microbial or other infections, congestive or ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPB nucleic acids can be used for the recombinant production of HSPB, for detecting HSPB in standard hybridisation and amplification assays (for diagnosis and monitoring), in gene therapy, as antisense, triplex-forming or ribozyme therapeutic, for detecting related sequences or genetic variations, and for chromosomal mapping. HSPB are also used to raise specific antibodies (Ab) and to screen for agonists and antagonists (potential therapeutic agents). Ab are used to diagnose, or monitor, HSPB-related diseases (in usual immunoassays), as therapeutic antagonists, in competitive drug screens, and for purification of HSPB from natural sources.
XX	Sequence 1923	Bp, 538 A, 439 C, 471 G, 475 T, 0 other,			
QY	1	AAGAATGCAATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGA	60		
DB	251	AAGATGGCATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGA	310		
QY	61	GATGTGCTAAGCAATCATCACTACTGCTTTTATGTTAAAGCCAGAACAGATCTAT	120		
DB	311	GATGTGCTAAGCAATCATCACTACTGCTTTTATGTTAAAGCCAGAACAGATCTAT	370		
QY	121	GAGCGATTGGACCTCTGGTGTATCTGTTGATACCCAGACTGAGTGGCTCAAGAACCTA	180		
DB	371	GAGCGATTGGACCTCTGGTGTATCTGTTGATACCCAGACTGAGTGGCTCAAGAACCTA	430		
QY	181	GAAGAAAGGCATCAAAATTTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCA	240		
DB	431	GAAGAAAGGCATCAAAATTTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCA	490		
QY	241	CTGAGAGCAGTGAGAAATACCCCACTGGGAGAGGGGAGAAATACGCTGTGATGTGAG	300		
DB	491	CTGAGAGCAGTGAGAAATACCCCACTGGGAGAGGGGAGAAATACGCTGTGATGTGAG	550		
QY	301	CCAGAATTCATAAGATATGCCATCTCTGGGCTCTTGGACAGCATTTGGACTCTCCAGAA	360		
DB	551	CCAGAATTCATAAGATATGCCATCTCTGGGCTCTTGGACAGCATTTGGACTCTCCAGAA	610		
QY	361	GGCATTTACAGAGAGATCTCTGGTGTGACTCTTTGATGAACTGACAGAGAGGGCTCA	420		
DB	611	GGCATTTACAGAGAGATCTCTGGTGTGACTCTTTGATGAACTGACAGAGAGGGCTCA	670		
QY	421	GAAGCAAGAGGAAAGATTGTTGTTTATTAACCAACCTTACATCACTCAAGACGAGT	480		
DB	671	GAAGCAAGAGGAAAGATTGTTGTTTATTAACCAACCTTACATCACTCAAGACGAGT	730		
QY	481	CAATACCGAAGCAGAGGGGGCGGTGAGAGCTGCAGAGGTGGGGCTTTGGCATCTCTCAT	540		
DB	731	CAATACCGAAGCAGAGGGGGCGGTGAGAGCTGCAGAGGTGGGGCTTTGGCATCTCTCAT	790		
QY	541	CGATCCGGGGCTCTCTTCCATCTACAGTCTCTCAACAGGATTTACAGGAATACAGAGAT	600		
DB	791	CGATCCGGGGCTCTCTTCCATCTACAGTCTCTCAACAGGATTTACAGGAATACAGAGAT	850		
QY	601	GGCGTGGCCAAATTCACAGCCTGTATTTACGTGAGAGATGCAGAAATGATGTCAAGA	660		


```

Db      851  GGCCTGCCCAAGATTCACAGAGCTGTATTAAGGTGAAGATGACGAATATGATGCAAGA 910
Qy      661  ATGGCTTCTCATGGATCAAAATTTGTCTATTGACGTAAAGATGGGGGCAAAAGACTACCA 720
Db      911  ATGGCTTCTCATGGATCAAAATTTGTCTATTGACGTAAAGATGGGGGCAAAAGACTACCA 970
Qy      721  GATACGTATCTCTCAACACTGTAGACAGATCACTGGAGCAAAATATCCAAACAGGTT 780
Db      971  GATACGTATCTCTCAACACTGTAGACAGATCACTGGAGCAAAATATCCAAACAGGTT 1030
Qy      781  GTACTGTGTCAGTGAACATCTGACAGCTGGAGATGTTGGGCAAGGTCATGATGATGCG 840
Db      1031  GTACTGTGTCAGTGAACATCTGACAGCTGGAGATGTTGGGCAAGGTCATGATGATGCG 1090
Qy      841  GGTGAGGCTTTATATATATGGAAGCACTCTCACTATTAAGAATCTTGGGCTGGTCCA 900
Db      1091  GGTGAGGCTTTATATATGGAAGCACTCTCACTATTAAGAATCTTGGGCTGGTCCA 1150
Qy      901  AAGAGACTCTGCGGCTGGTCTCTGACCTGACAGAGAAAGAGGTGAGTGGCTTC 960
Db      1151  AAGAGACTCTGCGGCTGGTCTCTGACCTGACAGAGAAAGAGGTGAGTGGCTTC 1210
Qy      961  CAGTATTATCAGTTACACAGAGTAATATTTTCCAACTACCTTGGTATGAGTCTGAC 1020
Db      1211  CAGTATTATCAGTTACACAGAGTAATATTTTCCAACTACCTTGGTATGAGTCTGAC 1270
Qy      1021  GCAGGAACCTTTTACCACTGCGGCTGCAATTCATCTGACATGAGAAAGCCAGGCTTC 1080
Db      1271  GCAGGAACCTTTTACCACTGCGGCTGCAATTCATCTGACATGAGAAAGCCAGGCTTC 1330
Qy      1081  ATGAGAGAGTTATGAGCTCTGACAGCCCTCAATATCATCTAGTCTGAGCCATGGA 1140
Db      1331  ATGAGAGAGTTATGAGCTCTGACAGCCCTCAATATCATCTAGTCTGAGCCATGGA 1390
Qy      1141  GAAGGAGACAGATCACTTTTGTGATCCAAAGCTGAGTGCCTGAGCCAGTCTTATGAT 1200
Db      1391  GAAGGAGACAGATCACTTTTGTGATCCAAAGCTGAGTGCCTGAGCCAGTCTTATGAT 1450
Qy      1201  GACTTATACAGTATTTCTTCCATCATCACTCCACGAGACACACATCATCTATGAT 1260
Db      1451  GACTTATACAGTATTTCTTCCATCATCACTCCACGAGACACACATCATCTATGAT 1510
Qy      1261  CCAAGACGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTATGTTGTCAGAC 1320
Db      1511  CCAAGACGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTATGTTGTCAGAC 1570
Qy      1321  ATGGAAGAAATGCTGCTAGGTCC 1344
Db      1571  ATGGAAGAAATGCTGCTAGGTCC 1594

```

RESULT 12
ID ACCS0603 standard; cDNA; 2077 BP.
AC ACCS0603;

XX 12-JUN-2003 (first entry)
DE Human secreted protein coding sequence, SEQ ID 270.

XX Cardiac; antiarrhythmic; antiarteriosclerotic; vasotropic; cyrostatic;
KW vulnerable; antiinflammatory; nootropic; neuroprotective;
KW antiParkinsonian; gene therapy; human; cardiovascular disorder;
gene; ss.

OS Homo sapiens.

XX W0200295010-A2.

XX 28-NOV-2002.

XX

```

PF 19-MAR-2002; 2002WC-US09785.
XX
XX 21-MAR-2001; 2001US-277340P.
PR 19-JUN-2001; 2001US-306171P.
PR 13-NOV-2001; 2001US-331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX WPI; 2003-129429/12.
XX
XX Novel human secreted proteins, useful for detecting, preventing,
XX diagnosing, prognosticating, treating and/or ameliorating
XX cardiovascular disorders such as arrhythmia -
XX
XX Claim 21; SEQ ID 270; 1881bp; English.
XX
XX The present invention relates to novel human secreted proteins
XX (ABR4763-ABR48145) and their coding sequences (ACCS0344-ACCS0566). The
XX proteins and their coding sequences are useful for the preparation of a
XX diagnostic or pharmaceutical composition for diagnosing or treating a
XX cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
XX coronary arteriosclerosis, muscular disorders, reproductive disorders,
XX immune system disorders, pulmonary disorders, renal disorders,
XX gastrointestinal disorders, or cancerous diseases and conditions, for
XX proliferative disorders and/or cancerous diseases and conditions, for
XX wound healing and epithelial cell proliferation, to treat inflammation or
XX infection, for treating thrombosis and arteriosclerosis, for treating or
XX preventing neural damage which occurs in neuronal disorders or
XX neurodegenerative conditions such as Alzheimer's disease and Parkinson's
XX disease, to enhance bone and periodontal regeneration and aid in tissue
XX transplants or bone grafts, to prevent skin aging or hair loss, to
XX stimulate growth and differentiation of hematopoietic cells and bone
XX marrow cells when used in combination with other cytokines, to maintain
XX organs before transplantation or for supporting cell culture of primary
XX tissues, to increase or decrease differentiation or proliferation of
XX embryonic stem cells, or to modulate mammalian characteristics or
XX metabolism.
XX
XX Note: The sequence data for this patent was published in electronic
XX format and is available from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;

```

Query Match 99.8%; Score 1340.8; DB 25; Length 2077;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 AAGATGGCATCTCTAAGAGACCTTTGAAGAATAAAGAAATAGCCAGCTGGA 60
Db 204 AAGATGGCATCTCTAAGAGACCTTTGAAGAATAAAGAAATAGCCAGCTGGA 263
Qy 61 GATGTTGCTAAGCAATCAATCACTAGCTGTTATGTTAAGCCGAAACAGATCTAT 120
Db 264 GATGTTGCTAAGCAATCAATCACTAGCTGTTATGTTAAGCCGAAACAGATCTAT 323
Qy 121 GAGGATTTGGCACTTCTGTTGATCTGTTGACCCGAGATGAGTGGCTCCAGAACTTA 180
Db 324 GAGGATTTGGCACTTCTGTTGATCTGTTGACCCGAGATGAGTGGCTCCAGAACTTA 383
Qy 181 GAAAAAGCATCCAAATTATGTACAAAACCTGACAGACAGATGAGTGGCTGAGAAATTCAC 240
Db 384 GAAAAAGCATCCAAATTATGTACAAAACCTGACAGACAGATGAGTGGCTGAGAAATTCAC 443
Qy 241 CTGAGCCAGTGAAGATACCCCACTGGAGAGGGGAGAGAAATCACTGTTGATGCTGAG 300
Db 444 CTGAGCCAGTGAAGATACCCCACTGGAGAGGGGAGAGAAATCACTGTTGATGCTGAG 503
Qy 301 CCAAGATTTATTAAGATAGCATCTCTGAGTCTGAGCAGCATTTGGGACTCTCCAGAA 360
Db 504 CCAAGATTTATTAAGATAGCATCTCTGAGTCTGAGCAGCATTTGGGACTCTCCAGAA 563

```

```

QY 361 GGCATTACAGCAGAAAGTTCTGTGTGATCCTCTTTTCATGATGACAGAGAGGGCCTCA 420
DB 564 GGCATTACAGCAGAAAGTTCTGTGTGATCCTCTTTTCATGATGACAGAGAGGGCCTCA 623
QY 421 GAAGCAAGAGGGAAGATTTGTTTATTAACCACTTAACATCACTACTCAAGACGGTG 480
DB 624 GAAGCAAGAGGGAAGATTTGTTTATTAACCACTTAACATCACTACTCAAGACGGTG 683
QY 481 CAATACGGAAGCGAGGGGGGGGAGGAGCTGCGCAAGGGGGGGGCTTTGGCATCTCTCAT 540
DB 684 CAATACGGAAGCGAGGGGGGGGAGGAGCTGCGCAAGGGGGGGGCTTTGGCATCTCTCAT 743
QY 541 CGATCCGTCGCTCTCTCTCATCTAGATCTCTCAACAGGATTCAGAAATACAGGAT 600
DB 744 CGATCCGTCGCTCTCTCTCATCTAGATCTCTCAACAGGATTCAGAAATACAGGAT 803
QY 601 GGGGTGCCCCAAATTCCAACAGCTGTATTAAGGTGAAGATGCAAAATGATGTCAAGA 660
DB 804 GGGGTGCCCCAAATTCCAACAGCTGTATTAAGGTGAAGATGCAAAATGATGTCAAGA 863
QY 661 ATGGCTCTCATGGGATCAAAATTTGATTCAGCTTAAGATGGGGGCAAAAGACTACCA 720
DB 864 ATGGCTCTCATGGGATCAAAATTTGATTCAGCTTAAGATGGGGGCAAAAGACTACCA 923
QY 721 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAATATCCAGAACAGGTT 780
DB 924 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAATATCCAGAACAGGTT 983
QY 781 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAATATCCAGAACAGGTT 840
DB 984 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAATATCCAGAACAGGTT 1043
QY 841 GGTGGAGCTTTTATATATAGGAGAGCACTCTCACTTATTAAGATCTTGGCTGCGTCCA 900
DB 1044 GGTGGAGCTTTTATATATAGGAGAGCACTCTCACTTATTAAGATCTTGGCTGCGTCCA 1103
QY 901 AAGAGACTCTGCGGCTGTGCTCTGACTGACAGAGCAAGGTGAGTGGCTTC 960
DB 1104 AAGAGACTCTGCGGCTGTGCTCTGACTGACAGAGCAAGGTGAGTGGCTTC 1163
QY 961 CAGTATTATCAGTTACCAAGATTAATTTTCCAACTACAGTCTGTGATGAGTCTGAC 1020
DB 1164 CAGTATTATCAGTTACCAAGATTAATTTTCCAACTACAGTCTGTGATGAGTCTGAC 1223
QY 1021 GCAGGAAGCTTTTATATATAGGAGAGCACTCTGAGCACTGAGAAAGGCGAGCCATC 1080
DB 1224 GCAGGAAGCTTTTATATATAGGAGAGCACTCTGAGCACTGAGAAAGGCGAGCCATC 1283
QY 1081 ATGAGAGAGGTTATGAGCTGCTGAGAGCCCTCAATATCACTCAAGTCTGAGCCATGGA 1140
DB 1284 ATGAGAGAGGTTATGAGCTGCTGAGAGCCCTCAATATCACTCAAGTCTGAGCCATGGA 1343
QY 1141 GAAGGAGAGAGATCAATCACTTTTGAATCCAAAGCTGAGTCTGAGAGCCATCTATGAT 1200
DB 1344 GAAGGAGAGAGATCAATCACTTTTGAATCCAAAGCTGAGTCTGAGAGCCATCTATGAT 1403
QY 1201 GACTTATTAACAAGTATTTCTTTCATATATCTCCACGAGACACCATGATGATGAT 1260
DB 1404 GACTTATTAACAAGTATTTCTTTCATATCACTCCACGAGACACCATGATGATGAT 1463
QY 1261 CCAAGAGAGATGAATGTTGCTGCTGCTGTTGGGGGTGTTTCTTATATGTTGGACAGAC 1320
DB 1464 CCAAGAGAGATGAATGTTGCTGCTGCTGTTGGGGGTGTTTCTTATATGTTGGACAGAC 1523
QY 1321 ATGGAAGAAATGCTGCTAGGTCC 1344
DB 1524 ATGGAAGAAATGCTGCTAGGTCC 1547

```

RESULT 13
AB271331
ID AB271331 standard; cdna; 2077 BP.
XX

```

AC AB271331,
XX
XX
DT 04-APR-2003 (first entry)
XX
DE Secreted protein-encoding gene 142 cDNA clone HRA035, SEQ ID NO:152.
XX
XX Human, secreted protein; digestive disorder; gastrointestinal disorder;
XX mouth; oesophagus; stomach; small intestine; large intestine; liver;
XX biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
XX immune disorder; inflammation; infection; wound healing; drug screening;
XX chromosome identification; chromosome mapping; cytostatic; gene therapy;
XX anti-inflammatory; immunosuppressive; vulnery; chromosome 8q22.2;
XX gene; ss.
XX
OS Homo sapiens.
XX
XX WO200276488-A1.
XX
XX 03-OCT-2002.
XX
XX
XX 19-MAR-2002; 2002MO-US08276.
XX
XX
XX 21-MAR-2001; 2001US-277340P.
XX
XX 19-JUL-2001; 2001US-306171P.
XX
XX 13-NOV-2001; 2001US-331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-022900/02.
XX
XX P-PSDB; ABR00152.
XX
XX
XX New human secreted proteins and nucleic acids, useful for detecting,
XX preventing, diagnosing, prognosticating, treating and/or ameliorating
XX e.g. gastrointestinal diseases and disorders, or cancers -
XX
XX Claim 21; Page 841-842; 1216pp; English.
XX
XX
XX AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
XX protein genes, and ABR00011-ABP00299 represent the proteins they encode.
XX AB271479-AB271540 represent human secreted protein genomic fragments. The
XX invention also encompasses antibodies specific for the secreted proteins,
XX the use of the secreted proteins in drug screening, and recombinant
XX vectors and host cells comprising a nucleic acid of the invention. The
XX secreted proteins, nucleic acids encoding them, antibodies or antibody
XX fragments specific for the secreted proteins, and modulators of protein
XX activity are useful for diagnosing, treating, ameliorating or preventing
XX digestive disorders. Such conditions include disorders of the mouth,
XX oesophagus, stomach, small intestine, large intestine, liver, biliary
XX tract and pancreas, and include cancers of these organs and tissues. The
XX secreted proteins and their nucleic acids may also be used in the
XX treatment of immune disorders, inflammation, infection,
XX hyperproliferative disorders, and to promote wound healing. Nucleic acids
XX of the invention may be used for chromosome identification, chromosome
XX mapping, in gene therapy, for identifying individuals from minute
XX biological samples, as hybridisation probes, and as molecular weight
XX markers. The present sequence represents a human secreted protein-
XX encoding cDNA clone of the invention.
XX
XX
XX Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;
XX
XX Query Match 99.8%; Score 1340.8; DB 25; Length 2077;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 AAGATGGCATCTCTTAAGAGACTTTGAAGAATAAAGAAATAGCAGCTGGA 60
DB 204 AAGATGGCATCTCTTAAGAGACTTTGAAGAATAAAGAAATAGCAGCTGGA 263
QY 61 GATGTTCTAAAGCAATCACTAGCTGTTTATGTAAGCCAGAAACAGATCTAT 120
DB 264 GATGTTCTAAAGCAATCACTAGCTGTTTATGTAAGCCAGAAACAGATCTAT 323

```

QY 121 GAGGATGGGACTTCTGGTGTATCTGTTGGACCCAGACTGAGTGGCTCCAGAACTTA 180
 DB 324 GAGGATGGGACTTCTGGTGTATCTGTTGGACCCAGACTGAGTGGCTCCAGAACTTA 383
 QY 181 GAAAAAGCATTCGAAATTTATGTATCCAAACCTGACGAGAAATGGCTGAGAAAGTTGAC 240
 DB 384 GAAAAAGCATTCGAAATTTATGTATCCAAACCTGACGAGAAATGGCTGAGAAAGTTGAC 443
 QY 241 CTGAGCGCAGTGAAGTATCCCACTGGAGAGGGAGAGAAATGACGCTGTATGCTGGAG 300
 DB 444 CTGAGCGCAGTGAAGTATCCCACTGGAGAGGGAGAGAAATGACGCTGTATGCTGGAG 503
 QY 301 CCAAGAAATTCATTAAGATAGCATCTCTGGGCTTGGACACAGATTTGGGACTCTCCAGAA 360
 DB 504 CCAAGAAATTCATTAAGATAGCATCTCTGGGCTTGGACACAGATTTGGGACTCTCCAGAA 563
 QY 361 GGCATTACAGAGAAATTTCTGGTGTGACCTCTTTGATGAACTGACAGAAAGGCTTCA 420
 DB 564 GGCATTACAGAGAAATTTCTGGTGTGACCTCTTTGATGAACTGACAGAAAGGCTTCA 623
 QY 421 GAACCAAGAGGAAATTTGTTTATTAACCACTTACATCACTACTCAAGAGCGGTG 480
 DB 624 GAACCAAGAGGAAATTTGTTTATTAACCACTTACATCACTACTCAAGAGCGGTG 683
 QY 481 CAATACCGAAGCAGAGGGGGGCTGGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTT 540
 DB 684 CAATACCGAAGCAGAGGGGGGCTGGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTT 743
 QY 541 CGATCCGTCCTCTCTCATCTCAAGTCCTCAACAGGATTTACAGAAATACAGGAT 600
 DB 744 CGATCCGTCCTCTCTCATCTCAAGTCCTCTCAACAGGATTTACAGAAATACAGGAT 803
 QY 601 GGCGTCCCAAAATTTCCAAAGCCCTGTATTAACGGTGAAGATTCAGAAATATATGTCAGA 660
 DB 804 GGCGTCCCAAAATTTCCAAAGCCCTGTATTAACGGTGAAGATTCAGAAATATATGTCAGA 863
 QY 661 ATGCTTCTCATGGGATCAAAATTTGTCAATTCAGTCAAGATGGGGGCAAAAGCTTACCCA 720
 DB 864 ATGCTTCTCATGGGATCAAAATTTGTCAATTCAGTCAAGATGGGGGCAAAAGCTTACCCA 923
 QY 721 GATACGTATCTCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAACGGTT 780
 DB 924 GATACGTATCTCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAACGGTT 983
 QY 781 GTACTGTCAGTGAACATCTGACAGCTGGGATGTTGGGCAAGGTCATGATGATGATGTC 840
 DB 984 GTACTGTCAGTGAACATCTGACAGCTGGGATGTTGGGCAAGGTCATGATGATGATGTC 1043
 QY 841 GGTGGAGCTTTATATCATGAGGAAAGCACTCTCACTTATTAAGATCTTGGGCTGGTCCA 900
 DB 1044 GGTGGAGCTTTATATCATGAGGAAAGCACTCTCACTTATTAAGATCTTGGGCTGGTCCA 1103
 QY 901 AAGAGACTCTGCGGCTGGTGTCTGACTGACAGAGAAACAAAGTGAATGATGCTTTC 960
 DB 1104 AAGAGACTCTGCGGCTGGTGTCTGACTGACAGAGAAACAAAGTGAATGATGCTTTC 1163
 QY 961 CAGTATATCATGTTACAGAAAGTAAATTTTCCAACTACAGTCTGGTATGAGAGCTGAC 1020
 DB 1164 CAGTATATCATGTTACAGAAAGTAAATTTTCCAACTACAGTCTGGTATGAGAGCTGAC 1223
 QY 1021 GCAGGAACCTTTACCACTGGGCTGGCAATTCATGSCAGTGAAGAAAGCCAGGGCCATC 1080
 DB 1224 GCAGGAACCTTTACCACTGGGCTGGCAATTCATGSCAGTGAAGAAAGCCAGGGCCATC 1283
 QY 1081 ATGAGAGAGTTATGAGCTCTGACGCCCTCATATATCACTCAAGTCTTGAGCCATGGA 1140
 DB 1284 ATGAGAGAGTTATGAGCTCTGACGCCCTCATATATCACTCAAGTCTTGAGCCATGGA 1343
 QY 1141 GAAAGGACAGACATCACTTTTGGATCCAAGCTGAGTGGCTGGAGCCAGTCTACTTGAT 1200
 DB 1344 GAAAGGACAGACATCACTTTTGGATCCAAGCTGAGTGGCTGGAGCCAGTCTACTTGAT 1403

QY 1201 GACTTATACAGATATTTCTTCTTCATCACTCTCCACGAGACACATGATGTCATGAT 1260
 DB 1404 GACTTATACAGATATTTCTTCTTCATCACTCTCCACGAGACACATGATGTCATGAT 1463
 QY 1261 CCAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTTGGCTGTTGTTCTTATGTTTGCAGAC 1320
 DB 1464 CCAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTTGGCTGTTGTTCTTATGTTTGCAGAC 1523
 QY 1321 ATGGAAGAAATGCTGCCCTAGGTCC 1344
 DB 1524 ATGGAAGAAATGCTGCCCTAGGTCC 1547
 RESULT 14
 ID AAK94491 standard; cDNA; 1860 BP.
 AC AAK94491;
 DT 06-NOV-2001 (first entry)
 XX Human full-length cDNA, SEQ ID NO: 3328.
 DB Human full-length cDNA; cDNA synthesis; oligo-capping; ss.
 XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX Homo sapiens.
 XX EPI130094-A2.
 XX 05-SEP-2001.
 PD 07-JUL-2000; 2000EP-0114089.
 PF 08-JUL-1999; 99JP-0194486.
 ER 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y,
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR P-PSDB; AAM93559.
 XX 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 PS Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a full length
 CC human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 SQ Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;
 Query Match 99.4%; Score 1336; DB 22; Length 1860;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1339; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 AAGATGGCATCTCTTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGGA 60
 DB 201 AAGATGGCATCTCTTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGGA 260

QY	61	GATGTGGTAAAGCAATCATCAACCTACTGTGTTATGGTAAAGCCCAAGACATCCAT	120
Db	261	GATGTGCTAAAGCAATCATCAACCTACTGTGTTATGGTAAAGCCCAAGACATCCAT	320
QY	121	GAGCGATTGGCACTTCTGGTTGATTAAGTGTGGAACCCACAAGTGGCTCCAAAGACCTA	180
Db	321	GAGCGATTGGCACTTCTGGTTGATTAAGTGTGGAACCCACAAGTGGCTCCAAAGACCTA	380
QY	181	GAATAAGCATCCAAATTATGTACCAAACTGCAAGCAAGTGGCTGGAGAAAGTTAC	240
Db	381	GAATAAGCATCCAAATTATGTACCAAACTGCAAGCAAGTGGCTGGAGAAAGTTAC	440
QY	241	CTGGAGCCAGTGAAGAAATACCCACCTGGGAGAGGGGAGAAAGAAATCAAGCTGATAGTGGAG	300
Db	441	CTGGAGCCAGTGAAGAAATACCCACCTGGGAGAGGGGAGAAAGAAATCAAGCTGATAGTGGAG	500
QY	301	CCAAGAAATTCAAGATAGCCATCCGTGGGTCTTGGACACAGCATTTGGGACTCTCCAGAA	360
Db	501	CCAAGAAATTCAAGATAGCCATCCGTGGGTCTTGGACACAGCATTTGGGACTCTCCAGAA	560
QY	361	GGCATTTACAGAGAAAGTTCTGGTGGTGAACCTCTTTTGATGAACCTGCAGAGAAAGGCTCA	420
Db	561	GGCATTTACAGAGAAAGTTCTGGTGGTGAACCTCTTTTGATGAACCTGCAGAGAAAGGCTCA	620
QY	421	GAACCAAGAGGGAAGATGTTGTTTATTAACAACCTTACATCACTACTCAAGAGCGGTG	480
Db	621	GAACCAAGAGGGAAGATGTTGTTTATTAACAACCTTACATCACTACTCAAGAGCGGTG	680
QY	481	CAATACCGAAAGCAAGGGGGCGGTGGAGCTGCAAGGTGGGGGCTTTGGCAATCTCTCAAT	540
Db	681	CAATACCGAAAGCAAGGGGGCGGTGGAGCTGCAAGGTGGGGGCTTTGGCAATCTCTCAAT	740
QY	541	CGATCCGTGGGCTCTCTTCATCTACAGTCCCTCACAGAGTATTCAAGAAATACCAAGAT	600
Db	741	CGATCCGTGGGCTCTCTTCATCTACAGTCCCTCACAGAGTATTCAAGAAATACCAAGAT	800
QY	601	GGCGGTGCCCAAAATTTCCAAACAGCCGTGATTAACGGTGAAGAAAGATGCAGAAATGATGTCAAGA	660
Db	801	GGCGGTGCCCAAAATTTCCAAACAGCCGTGATTAACGGTGAAGAAAGATGATGCAGAAATGATGTCAAGA	860
QY	661	ATGGCTTCTCATGGGATCAAAATTGTCAATTAGCTGAAAGATGGGGGGCAAGACCTACCCA	720
Db	861	ATGGCTTCTCATGGGATCAAAATTGTCAATTAGCTGAAAGATGGGGGGCAAGACCTACCCA	920
QY	721	GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACGGTT	780
Db	921	GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACGGTT	980
QY	781	GTACTCTGTCAGTGAACATCTGGACAGCTGGGAATGTTGGGACAGGGTCCATGGATGATGGC	840
Db	981	GTACTCTGTCAGTGAACATCTGGACAGCTGGGAATGTTGGGACAGGGTCCATGGATGATGGC	1040
QY	841	GGTGGAGCCTTTATATCAATGGGGAAGCACTCACTAATTAAGATCTTGGGCTGGTCCCA	900
Db	1041	GGTGGAGCCTTTATATCAATGGGGAAGCACTCACTAATTAAGATCTTGGGCTGGTCCCA	1100
QY	901	AAGAGGACTCTGCGGCTGCTCTGCACTGCAGAGAAACAAGGTGAGTGTGTCCTTC	960
Db	1101	AAGAGGACTCTGCGGCTGCTCTGCACTGCAGAGAAACAAGGTGAGTGTGTCCTTC	1160
QY	961	CAGTATTATCACTTACACAGAGTAATATTTTCCAACTACAGTCTGTGATGGAGTCTGAC	1020
Db	1161	CAGTATTATCACTTACACAGAGTAATATTTTCCAACTACAGTCTGTGATGGAGTCTGAC	1220
QY	1021	GCAGGAACCTTCTTACCCACTGGGCTGGAATTCACATGGCAGTGAAGAAAGGCAAGGGCCATC	1080
Db	1221	GCAGGAACCTTCTTACCCACTGGGCTGGAATTCACATGGCAGTGAAGAAAGGCAAGGGCCATC	1280
QY	1081	ATGGAAGAGATTATAGACCTGCTGCAGGCCCTCAATATCACTCAAGGTCTTAGCCATAGGA	1140
Db	1281	ATGGAAGAGATTATAGACCTGCTGCAGGCCCTCAATATCACTCAAGGTCTTAGCCATAGGA	1340
QY	1141	GAAAGGACAGACATCACTTTTGATTCAGAGCTGGAGTGGCTTGAGCAAGTCTACCTTGAT	1200

Db	1341	GAGGGGACAGACATCACTTTTGATTCAGAGTGCTGGAGCAGCTCACTTGAAT	1400
Qy	1201	GACTTATACAGATATTTCTTCTTCATCACTCTCCAGGAGACACCAATGACTGCATGGAT	1260
Db	1401	GACTTATACAGATATTTCTTCTTCATCACTCTCCAGGAGACACCAATGACTGCATGGAT	1460
Qy	1261	CCAAAGCAGATGAATGTTGCTGCTGCTGTTGGGCGTGTGTTCTTATGTTGTCAGAC	1320
Db	1461	CCAAAGCAGATGAATGTTGCTGCTGCTGTTGGGCGTGTGTTCTTATGTTGTCAGAC	1520
Qy	1321	ATGGAGAAATGCGCTAGCTCC	1344
Db	1521	ATGGAGAAATGCGCTAGCTCC	1544
RESULT 15			
AAH9703			
ID	AAH9703 standard; cDNA; 1895 BP.		
XX	AAH9703;		
AC			
DT	16-OCT-2001 (first entry)		
XX			
DE	Human protein encoding cDNA sequence SEQ ID NO:538.		
XX			
XX	Human; cancer; ulcer; HIV infection; human immunodeficiency virus;		
KW	antiflammatory; antirheumatic; antiarthritic; immunosuppressive;		
KW	antibacterial; endocrine; cardiac; central nervous system; virucide;		
KW	anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;		
KW	antigregant; haemostatic; vulnery; antilucer; osteopathic; eczema;		
KW	dermatological; antiallergic; antiaesthetic; antidiabetic; cystostatic;		
KW	neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;		
KW	immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;		
KW	antimaphilactic; rheumatoid arthritis; septic shock; pancreatitis;		
KW	cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;		
KW	genetic disease; haematopoietic disorder; platelet disorder; asthma;		
KW	thrombocytopaenia; osteoporosis; severe combined immunodeficiency;		
KW	allergic rhinitis; diabetes; multiple sclerosis; depression;		
KW	Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;		
KW	neurological disorder; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	MO2001S3455-A2.		
XX			
PD	26-JUL-2001.		
XX			
XX			
PF	22-DEC-2000; 2000WC-US35017.		
XX			
XX			
PR	23-DEC-1999; 99US-0471275.		
XX			
PR	21-JAN-2000; 2000US-0488725.		
XX			
PR	25-APR-2000; 2000US-0552317.		
XX			
PA	(HYSE-) HXSEQ INC.		
XX			
PI	Tang YT, Liu C, Drmanac RT;		
XX			
DR	WPI: 2001-457603/49.		
XX			
XX	P-PSDB; AAM25762.		
PT	Isolated human polynucleotides encoding polypeptides, useful for the		
XX	treatment and diagnosis of e.g. cancer, ulcers and HIV infection -		
XX			
P8	Claim 1; Page 591; 1217pp; English.		
XX			
CC	AAH9166 to AAH9904 encode the human proteins given in AAM25225 to		
CC	AAM2563. The proteins can have activities based on the tissues and		
CC	cells they are expressed in, such as: antinflammatory; antirheumatic;		
CC	antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;		
CC	central nervous system; virucide; anti-HIV; fungicide; antimutagen;		
CC	cardiovascular; antianaemic; antigregant; haemostatic; vulnery;		
CC	antilucer; osteopathic; dermatological; antiallergic; antiaesthetic;		

CC antidiabetic; cyostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.

XX Sequence 1895 BP; 530 A; 439 C; 450 G; 476 T; 0 other;

Query Match 95.5%; Score 1283.6; DB 22; Length 1895;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1335; Conservative 0; Mismatches 9; Indels 6; Gaps 4;

QY 1 AAGATGGCATCTCTAAGAGGCTTTTGAAGAAATGAAGAAATGCCAGCTGTGA 60
 DB 198 AAGATGGCATCTCTAAGAGGCTTTTGAAGAAATGAAGAAATGCCAGCTGTGA 257
 QY 61 GATGTTGCTAAAGCAATCATCACTAGCTGTTTAAAGCCAGAACAGATCTAT 120
 DB 258 GATGTTGCTAAAGCAATCATCACTAGCTGTTTAAAGCCAGAACAGATCTAT 317
 QY 121 GAGCGATTGGCACTTCTGTTGTAATCTGTGGACCCAGACTGAGTGTCCAGAACCTTA 180
 DB 318 GAGCGATTGGCACTTCTGTTGTAATCTGTGGACCCAGACTGAGTGTCCAGAACCTTA 377
 QY 181 GAAAGAGCATCTCAATATATGTAACCAAACTGCGAGCAAGTGGCTGTGAAGAAATTAC 240
 DB 378 GAAAGAGCATCTCAATATATGTAACCAAACTGCGAGCAAGTGGCTGTGAAGAAATTAC 437
 QY 241 CTGAGCCAGTGAATACCCCACTGAGAGGAGGAGAAATCACTGTGATGTGGAG 300
 DB 438 CTGAGCCAGTGAATACCCCACTGAGAGGAGGAGAAATCACTGTGATGTGGAG 497
 QY 301 CCAGAAATTCATAGATAGCCATCTCGGCTCTGGCAGCAGCATTTGGAGCTCTCCAGAA 360
 DB 498 CCAGAAATTCATAGATAGCCATCTCGGCTCTGGCAGCAGCATTTGGAGCTCTCCAGAA 557
 QY 361 GGCATTACAGCAGAGTCTGTTGTAATCTGTGGACCTTCTGTAAGTGAAGAAAGGCTTCA 420
 DB 558 GGCATTACAGCAGAGTCTGTTGTAATCTGTGGACCTTCTGTAAGTGAAGAAAGGCTTCA 617
 QY 421 GAAGCAAGAGGAGATGTTGTTTATACCAACTTACATCACTACATCAAGACGGTG 480
 DB 618 GAAGCAAGAGGAGATGTTGTTTATACCAACTTACATCACTACATCAAGACGGTG 677
 QY 481 CAATACCAAGCAGCAGGAGGCTGGAAGCTGCAAGGTGGGGCTTTGGCATCTCTCAT 540
 DB 678 CAATACCAAGCAGCAGGAGGCTGGAAGCTGCAAGGTGGGGCTTTGGCATCTCTCAT 737
 QY 541 CGATCGGAGCTCTCTTCTCACTCAAGTCTCTCAAGATTCAGAGAAATACAGAGAT 600
 DB 738 CGATCGGAGCTCTCTTCTCACTCAAGTCTCTCAAGATTCAGAGAAATACAGAGAT 797
 QY 601 GGCCTGCCAAATTCAGACAGCTGTATACGTTGAAGATGAGAAATGATGTCAAGA 660
 DB 798 GGCCTGCCAAATTCAGACAGCTGTATACGTTGAAGATGAGAAATGATGTCAAGA 857
 QY 661 ATGGCTTCTCATGGGATCAAAATGTCTATTCAGCTAAAGATGGGGCAAGACCTACCA 720
 DB 858 ATGGCTTCTCATGGGATCAAAATGTCTATTCAGCTAAAGATGGGGCAAGACCTACCA 917
 QY 721 GATCTGATCTCTTCAACACTGTAGCAGATCACTGGAGCAAAATATCCAGAACAGTT 780
 DB 918 GATCTGATCTCTTCAACACTGTAGCAGATCACTGGAGCAAAATATCCAGAACAGTT 977

QY 781 GTACTGTCAGTGAGCATCTGAGCAGCTGGAGATTTGGCAGGGTGCATGATGATGC 840
 DB 978 GTACTGTCAGTGAGCATCTGAGCAGCTGGAGATTTGGCAGGGTGCATGATGATGC 1037
 QY 841 GGTGAGCCTTTATATATCATGAGGAAGCACTTCACTTATTAAGATCTTGGCTGCTCA 900
 DB 1038 GGTGAGCCTTTATATATCATGAGGAAGCACTTCACTTATTAAGATCTTGGCTGCTCA 1097
 QY 901 AAGGAGCTCTGCGGCTGGTCTCTGAGCTGCAAGAAACAAGGTGAGTGTGCTTC 960
 DB 1098 AAGGAGCTCTGCGGCTGGTCTCTGAGCTGCAAGAAACAAGGTGAGTGTGCTTC 1157
 QY 961 CAGATTTATCAGTTACACAAGATTAATATTTCCAACTACAGTCTGATGATGATGAC 1020
 DB 1158 CAGATTTATCAGTTACACAAGATTAATATTTCCAACTACAGTCTGATGATGATGAC 1217
 QY 1021 GCAGAACTTTCTTACCCACTGAGCTGCAATTCATGAGAGTGAAGAAAGCCAGGCTATC 1080
 DB 1218 GCAGAACTTTCTTACCCACTGAGCTGCAATTCATGAGAGTGAAGAAAGCCAGGCTATC 1277
 QY 1081 ATGAGAGGTTATGAGCCTGCTGCAAGCTCCCAATTCATGAGTCTGAGCCTATGA 1140
 DB 1278 ATGAGAGGTTATGAGCCTGCTGCAAGCTCCCAATTCATGAGTCTGAGCCTATGA 1337
 QY 1141 GAAGGACAGACATCACTTTTGGATCCAGCTGAGAGTGTGCTGAGCAGTCTTATGAT 1200
 DB 1338 GAAGGACAGACATCACTTTTGGATCCAGCTGAGAGTGTGCTGAGCAGTCTTATGAT 1397
 QY 1201 GACTTATACAGATATTTCTTCTTCATCACTCCCAAGAGACACATGACTGTC--ATG 1257
 DB 1398 GACTTATACAGATATTTCTTCTTCATCACTCCCAAGAGACACATGACTGTCATGGG 1457
 QY 1258 GATCCAAAGAGATGATG--TTGCTGCTGCTG--TTTGGGCTGTTTCTTATGT--TGT 1314
 DB 1458 ATCCAAAGAGATGATGTTTCTGCTGCTGTTTGGGCTGTTTCTTATGTGTGT 1517
 QY 1315 GCAGACATGGAAGAAATGCTGCTAGGTCC 1344
 DB 1518 GCAGACATGGAAGAAATGCTGCTAGGTCC 1547

Search completed: December 22, 2003, 15:10:06
 Job time : 328.93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:22:34 ; Search time 2630.64 Seconds
(without alignments)
12417.225 Million cell updates/sec

Title: US-09-745-763-35_COPY_171_1514
Perfect score: 1344
Sequence: 1 AAGAAATGCTCTTAAGAG.....AAGAAATGCTGCTAGGTCC 1344

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vtl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query length	ID	Description
1	1088	81.0	1868	11 BC010977 Mus muscu
2	1086.4	80.8	1739	11 AK032972 Mus muscu
3	1084.8	80.7	1805	11 AK075686 Mus muscu
4	886.8	66.0	1024	11 BC012019 Homo sapi

5	833.8	62.0	1050	13 BX439467	BX439467
6	771.8	57.4	1038	13 BX355940	BX355940
7	740.6	55.1	866	13 BU146905	BU146905
8	740	55.1	902	10 BG751497	BG751497
9	726.6	54.1	996	13 BX355939	BX355939
10	724.8	53.9	909	13 BQ878966	BQ878966
11	719.2	53.5	739	14 CB958693	CB958693
12	702.4	52.3	1126	13 BX360507	BX360507
13	694.2	51.7	919	13 BX355995	BX355995
14	682.4	50.8	813	14 CD519131	CD519131
15	681.8	50.7	708	14 CD366462	CD366462
16	672.4	50.0	1012	13 BX416896	BX416896
17	664.6	49.4	998	13 BX460463	BX460463
18	653.2	48.6	1020	13 BX439466	BX439466
19	636.8	47.4	653	10 BE218907	BE218907
20	630.4	46.9	650	10 BG432644	BG432644
21	626.2	46.6	804	14 CD352906	CD352906
22	625.6	46.5	677	10 BE670671	BE670671
23	614.8	45.7	1073	13 BX360506	BX360506
24	612.2	45.6	889	10 BE906771	BE906771
25	611.4	45.5	922	12 BG761741	BG761741
26	604.6	45.0	945	12 BI909780	BI909780
27	604	44.9	926	13 BQ938234	BQ938234
28	599.6	44.6	805	10 BG563740	BG563740
29	597.4	44.4	3151	11 BC017373	BC017373
30	594.8	44.3	1081	13 BQ072892	BQ072892
31	591.6	44.0	1081	12 BI754143	BI754143
32	591.2	44.0	1003	13 BX416895	BX416895
33	583.8	43.4	701	12 BI771109	BI771109
34	576.8	42.9	751	14 BC558931	BC558931
35	570.4	42.4	786	10 BG430966	BG430966
36	569.8	42.4	783	12 BI754468	BI754468
37	568.4	42.3	929	13 BX335994	BX335994
38	560.8	41.7	822	12 BI911772	BI911772
39	560	41.7	944	12 BI819273	BI819273
40	554.2	41.0	879	13 BQ218838	BQ218838
41	551	41.0	937	14 CD515740	CD515740
42	542.4	40.4	544	10 AW952474	AW952474
43	534.6	39.8	794	14 CD356404	CD356404
44	534.4	39.8	774	14 CB308109	CB308109
45	530.2	39.4	727	12 BG765433	BG765433

ALIGNMENTS

RESULT 1
BC010977
LOCUS
DEFINITION Mus musculus, similar to plasma glutamate carboxypeptidase, clone
IMAGE:4018296, mRNA.
ACCESSION BC010977
VERSION BC010977.1 GI:15012137
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbioology.org>

contact: amadan@systembiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM, at: <http://image.llnl.gov>
Series: IRAX Plate: 18 Row: 9 Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9055233
This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers

1..1868
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH 11"
/db_xref="taxon:10090"
/clone="IMAGE:4018296"
/issue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."
/clone_1id="NCI CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 505 a 418 c 460 g 485 t
ORIGIN

Query Match 81.0%; Score 1088; DB 11; Length 1868;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 1184; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

1 AAGAATGCGATCTCTAAGAGAGACTTTGAGAAATTAAGAAATATGAGCCAGCTGTGA 60
267 AAGATGCTGTTTTCACGCGAATTTTGAGAAATTAAGAAATATGAGCCAGCTGTGA 326
61 GATGTTGCTAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAGACAGATCTTAT 120
327 GATGTTGCTAAGCAATCATCACTAGCTGTTTATGTTAAATACAGACCGGCTTAT 386
121 GAGCATTTGCACTTCTGTTGATATCTGTGGAACCCAGATGAGTCTCCAGAACCTA 180
387 GAGCGTTTGGGACTTCTAGTTGATATCTGTGGAACCCAGATGAGTCTCCAGAACCTA 446
181 GAAAAAGCCATTCATATATGTAACCAAACTGACAGAAATGAGCTGAGAAATTTTAC 240
447 GAGAAAGCTATTCATATATGTAACCAAACTGACAGAAATGAGCTGAGAAATTTTAC 506
241 CTGAGCAGTGAAGATACCCCACTGAGAGAGAGAGAAATGAGTGTGATGTGAG 300
507 CTGAGCAGTGAAGATACCCCACTGAGAGAGAGAGAAATGAGTGTGATGTGAG 566
301 CCAAGATTCATATATGATACCTCTGAGTCTTGGACAGATTTGGATCTCTCCAGAA 360
567 CCGAATTCATATATGATACCTCTGAGTCTTGGACAGATTTGGATCTCTCCAGAA 626
361 GGCATTACAGAGAGTCTGTTGTTGATCTCTTTCATGTAATCTGACAGAGAGGCTTCA 420
627 GGCATCAAGAGAGTCTGTTGTTGATCTCTTTCATGTAATCTGACAGAGAGGCTTCA 686
421 GAAGCAGAGAGAGATGTTGTTTATACCAACCTTACATCACTACTCAAGAGCGGTG 480
687 GAAGCAGAGAGAGATGTTGTTTATACCAACCTTACATCACTACTCAAGAGCGGTG 746
481 CAATTCAGAGAGAGCGGCGGTGAGAGTGTGCAAGTGTGGGGCTTTGGCATCTTCA 540
747 CAGTACCGGTTGCAAGAGCTGTGAGAGCTGTGCAAGTGTGGGGCTTTGGCATCTTCA 806
541 CGATCGTGGCGCTCTTCCATCTACAGTCTCTCAAGAGATTCAGAGATTCAGAGAT 600
807 CAATGATGAGCTCTCTTTCATCTACAGTCTCTCAAGAGATTCAGAGATTCAGAGAT 866
601 GGCCTGCCCAAAATTCAGAGAGCTGTATTCAGTGTGAGAGATGAGAAATGATGTCA 660
867 GGTGTGCCCAAGATTCAGAGAGCTGTATTCAGTGTGAGAGATGAGAAATGATGTCTGA 926

661 ATGCGTCTCATGGGATCAAAATTTGATTCAGCTAAAGATGGGGGCAAGAGACTTACCA 720
927 ATGGCTTCTGTGGGAAAGAAATTTGATTCATTCATTTGAAATGGAGCAAGACTTATCA 986
721 GATACGATTCCTTCAACACTGTGAGAGATCACTGGGAGCAAAATTCAGAACAGATT 780
987 GATACGATTCCTTCAACACTGTGAGAGATCACTGGGAGCAAAATTCAGAACAGATT 1046
781 GTACTGTGAGTGAACATCTGAGAGCTGGAGATTTGGGAGGATTCAGATGATGGC 840
1047 GTCTGTGAGTGAACATCTGAGAGCTGGAGATTTGGGAGGATTCAGATGATGGC 1106
841 GGTGAGCGCTTTATATATGAGGAAAGCACTCTCACTTATTAAGATTTGGGCTCGTCCA 900
1107 GTTGAAGCTTTCATATATGAGGAGAGCACTCTCACTTATTAAGATTTGGGCTCGTCCA 1166
901 AAGAGACTCTGCGGCTGTGCTCTGAGCTGACAGAGAGCAAGGTGAGTGTGCTTC 960
1167 AAGAGACTCTGCGGCTGTGCTCTGAGCTGACAGAGAGCAAGGTGAGTGTGCTTC 1226
961 CAGTATTCATGATTCACAAAGTAAATTTTCCATCACTGCTGTGATGAGTCTGAC 1020
1227 CAGTATTCATGATTCACAAAGTAAATTTTCCATCACTGCTGTGATGAGTCTGAC 1286
1021 GCAGGAACCTTCTTACCCAGCTGGGCTGCAATTCATCTGSCAGTGAAGGCGAGGCTATC 1080
1287 TCAGGAACCTTCTTACCCAGCTGGGCTGCAATTCATCTGSCAGTGAAGGCGAGGCTATC 1346
1081 ATGAGGAGGTTATGAGCTGCTGACAGCCCTCAATTCATCTGAGTCTGAGGATGGA 1140
1347 ATGAGGAGGTTATGAGCTGCTGACAGCCCTCAATTCATCTGAGTCTGAGGATGGA 1406
1141 GAAGGAGGATCATCACTTTTGGATCAAGCTGAGTGTGAGGCTGAGGCTGAT 1200
1407 GAAGGAGGATCATCACTTTTGGATCAAGCTGAGTGTGAGGCTGAGGCTGAT 1466
1201 GACTTATCAAGATTTTCTTCCATCACTCCACGAGAGACCAAGATGATGAT 1260
1467 GACTTATCAAGATTTTCTTCCATCACTCCACGAGAGACCAAGATGATGAT 1526
1261 CCAAGCAGATGATGTTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTTGGACAG 1320
1527 CCAAGCAGATGATGTTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTTGGACAG 1586
1321 ATGGAAGATGCTGCTGATGCTC 1344
1587 ATGATGAAATGTTGCCAGGCTCC 1610

RESULT 2
AK032972
LOCUS 1739 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus 12 days embryo male wolffian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:6720483C06
ACCESSION AK032972
VERSION AK032972.1 GI:26328732
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
9879253
PUBMED 10349636
2
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to

JOURNAL
MEDLINE
PUBMED
AUTHORS
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

JOURNAL
MEDLINE
PUBMED
AUTHORS
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, S., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kiehl, P., Lewis, S., Matsuo, Y., Nakai, I., Perole, G., Quackenbush, J., Schriml, L. M., Staahl, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bolt, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, D., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-Oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

JOURNAL
MEDLINE
PUBMED
AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1739)

JOURNAL
REFERENCE
AUTHORS
ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P., FUKUDA, S., FURUNO, M., HANAGAKI, T., HARA, A., HASHIZUME, W., HAYASHIDA, K., HAYASE, N., HIRAMOTO, K., HIROKA, T., HIROKANE, T., HORI, F., IMOCANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T., KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONO, H., KUNDA, M., KOYA, S., KUTAHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M., NAKAMURA, M., NISHI, K., NOMURA, K., NAMAZAKI, R., OHNO, M., OHSAITO, N., OKAZAKI, Y., SAITO, R., SAITO, H., SAKAI, C., SAKAI, K., SAKAZUME, N., SANO, H., SASEKI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T., SOGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-ANAHARA, S., TAKEDE, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A., MURAMATSU, M. and HAYASHIZAKI, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES		Please visit our web site for further details.	
source		URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers	
1..1739		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/strain="C57BL/6J"	
		/db_xref="FANTOM:DB:6720483C06"	
		/db_xref="taxon:10090"	
		/clone="6720483C06"	
		/sex="male"	
		/tissue type="olfactory duct includes surrounding region"	
		/clone lib="RIKEN full-length enriched mouse cDNA library"	
		/dev stage="12 days embryo"	
CDS		81..1493	
		/note="unnamed protein product; plasma glutamate carboxypeptidase (MGI:1889205, GI:1018755, evidence: BLASTN, 98%, match=943)"	
		putative"	
		/protein_id="BAC28105.1"	
		/db_xref="GI:26328733"	
		/translation="MRSLEFLFVHLALGSKAVKNGVORTPREIKEEIANVEDVAKAIIIMLAYKXONSYERLGLVDVTPRISGSRLKALDIOMYONQOQDELNVHLEQVRIPIHMERGESAVMEPRIRHKALIGLSSIGTPGGITAEVLVAVSPFELGRASBARAKIIVNPNPYGKTYVRQGVAAKVGAVASLIQSVASPSIYSPHGIQKYDDGPKPIPTACTIYEDAMSRMSRNKVIYIHEMGAKTYPDPSPTVAETIGSMYEEVLVLSNDVQGLDDGGAFVSEALSLVKDLSLRKRTLVLMWLEEOGGLASQYVEIHLKNDISKYSLVMDADGTFPLPTGLQFGASRAIKKEVNNLQPLNTAKTFSNGEGDINFWIOAGVPGASLRDLDTLYKFFHSHSDTWTMDPKQNNVAAAWVAAYVAVADMDENLPRS"	
		1722..1727	
		polya_signal	
		/note="putative"	
		polya_site	
		/note="putative"	
BASE COUNT		482 a 369 c 423 g 465 t	
ORIGIN			
Query Match		80.8%; Score 1086.4; DB 11; Length 1739;	
Best Local Similarity		88.0%; Pred. No. 0;	
Matches 1183; Conservative 0; Mismatches 161; Indels 0; Gaps 0;			
1	AAGATGGCATCTCTAAGAGACTTTGAAGAATAAAGAGAAATAGCAGCTGGA	60	
147	AAGATGGCTTTCTCGACCAATTTCGAAATATAAGAAATAGCACTATGAA	206	
61	GATGTGCTAAGCAATCATCAACCTAGCTGTTATAGTTAAAGCCCGAAGATCCTAT	120	
207	GATGTGCTAAGCAATCATCAACCTAGCTGTTATAGTTAAATACCAAGACCGTCTAT	266	
121	GAGCGATGGCACTCTGCTGTTGATCTGTTGAACCCAGACTGAGTGGCTCAAGAACTTA	180	
267	GAGCGTTGGAGCTTCTAGTTGATCTGTTGAGCCAGCTGAGTGGCTCTTAAGAACTTA	326	
181	GAAGAAGCATCAATTAATTAATGTCACAAACCTGACAGAAATAGGGCTGAGAAAGTTTAC	240	
327	GAGAAACCTTCAATCATGATCAAAACCTGACGAAATAGGGCTGAGAAAGTTTAC	386	
241	CTGAGGCACTGAGATATACCCCACTGGAGAGGAGAGAAATAGCTGATGCTGGAG	300	
387	CTGAGGCACTGAGATATACCCCACTGGAGAGGAGAGAAATAGCTGATGCTGGAG	446	
301	CGAAGATTAATAGATAGCATCTCGGCTCTTGGCAGCAGCATTTGGGACTCTCCAGAA	360	
447	CTTGAATTAATAGATAGCATCTCGGCTCTTGGCAGCAGCATTTGGGACTCTCCAGAA	506	
361	GGAATTAATAGATAGCATCTCGGCTCTTGGCAGCAGCATTTGGGACTCTCCAGAA	420	
507	GGAATTAATAGATAGCATCTCGGCTCTTGGCAGCAGCATTTGGGACTCTCCAGAA	566	
421	GAAGCAAGAGGAAATGTTGTTTAAACCACTTACATCACTACAGAGCAGCTG	480	

Db	567	GAAGCAAGAGGAGATCATTTGTTATTAACACAGCTTACATGCTATGAGAAAGACTGTG	626
Qy	481	CAATACCAAGCAGCGGGCGGTGAACTGCGCAAGGTGGGGGCTTTGGCATCTCTATT	540
Db	627	CAGTACCGGGTGCAGGAGCTGTGAACTGCGCAAGGTGGGGGCTTTGGCATCTCTATT	686
Qy	541	CGATCCGCGGCTCTCTCCATCTACAGTCTCTCACACAGGATTCAGAAATACCAAGAT	600
Db	687	CAATCAGTACGCTCTCTTCCATCTACAGTCTCTCACACAGGATTCAGAAATACCAAGAT	746
Qy	601	GCGCGTCCCAAAATTCACACAGCTGTATTAACGTTGAGAGATGCGAAATGATGTCAGA	660
Db	747	GGTGTGCCCAAGATTCACACAGCTGTATTAACGTTGAGAGATGCGAAATGATGTCAGA	806
Qy	661	ATGGTTCTCATGGGATCAAAATTTCTCATGCTCAAGATGGGGGCAAGACCTACCCA	720
Db	807	ATGGTTCTCATGGGATCAAAATTTCTCATGCTCAAGATGGGGGCAAGACCTACCCA	866
Qy	721	GATATCTGATTCCTTCAACATCTGTACAGATCACTGGGAGCAAAATATTCAGAACAGT	780
Db	867	GATATCTGATTCCTTCAACATCTGTGTACAGATCACTGGGAGCAAAATATTCAGAACAGT	926
Qy	781	GTAATGTCAGTGAACATCTGACAGCTGGAGATGTTGGGCAAGGTGCCATGATGATGAT	840
Db	927	GTCCTGTCAGTGAACATCTGACAGCTGGAGATGTTGGGCAAGGTGCCATGATGATGAT	986
Qy	841	GCTGAGGCTTTATATTCATGGAAGCACTCTCATTTAAGATCTTGGGCTGGCTTCA	900
Db	987	GCTGAGGCTTTATATTCATGGAAGCACTCTCATTTAAGATCTTGGGCTGGCTTCA	1046
Qy	901	AAGAGACTCTGCGGCTGCTGTGACCTGACAGTGAAGAAACAAGGTGATGCTCTTC	960
Db	1047	AAGAGACTCTGCGGCTGCTGTGACCTGACAGTGAAGAAACAAGGTGATGCTCTTC	1106
Qy	961	CAGTATTCAGATTAACACAGGTAAATATTTCCAACTACAGTCTGTGATGATGATGAT	1020
Db	1107	CAGTATTCAGATTAACACAGGTAAATATTTCCAACTACAGTCTGTGATGATGATGAT	1166
Qy	1021	GCAGGAACCTTTTCCCACTGGGCTGCAATTCATCTGGCAGTGAAGAAAGGCCAGCTTC	1080
Db	1167	GCAGGAACCTTTTCCCACTGGGCTGCAATTCATCTGGCAGTGAAGAAAGGCCAGCTTC	1226
Qy	1081	ATGGAGAGGATTAATGAGCTGCTGACGCGCTCAATATCATCTGAGCTGAGCCATGGA	1140
Db	1227	ATGGAGAGGATTAATGAGCTGCTGACGCGCTCAATATCATCTGAGCTGAGCCATGGA	1286
Qy	1141	GAAGGACAGACATCACTTTTGATTCACAGCTGAGTGCCTGAGCCAGTCTTACTTAT	1200
Db	1287	GAAGGACAGACATCACTTTTGATTCACAGCTGAGTGCCTGAGCCAGTCTTACTTAT	1346
Qy	1201	GACTTATTCAGATTAATGAGCTGCTGACGCGCTCAATATCATCTGAGCTGAGCCATGGA	1260
Db	1347	GACTTATTCAGATTAATGAGCTGCTGACGCGCTCAATATCATCTGAGCTGAGCCATGGA	1406
Qy	1261	CCAAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC	1320
Db	1407	CCAAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC	1466
Qy	1321	ATGGAAGAAATCTGCTAGGTGC 1344	
Db	1467	ATGGAAGAAATCTGCTAGGTGC 1490	

RESULT 3
AK075686
LOCUS
DEFINITION
AK075686 1805 bp mRNA linear HTC 07-DEC-2002
Mus musculus 18-day embryo whole body cDNA, Riken full-length
enriched library, clone:1190003P12 product:plasma glutamate
carboxypeptidase, full insert sequence.

ACCESSION
AK075686
VERSION
AK075686.1 GI:26344480
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohta, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, D., Boffelli, D., Boljunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mondet, P., Noridone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyono, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	1217851
REFERENCE	5
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	1217851

Db 1533 ATGATGAATGCTGCCAGGTCC 1556

|||||

RESULT 4
BC012019
LOCUS
DEFINITION Homo sapiens, similar to plasma glutamate carboxypeptidase, clone
IMAGE:4455631, mRNA.
ACCESSION EC012019.1 GI:15080558
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Strausberg, R.
TITLE 1 (bases 1 to 1024)
JOURNAL Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Louised, H.,
Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

FEATURES
Source
1..1024
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4455631"
/cissue_type="Kidney, hypernephroma"
/clone_id="NIH_MGC_89"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 317 a 210 c 256 g 241 t
ORIGIN

Query Match 66.0%; Score 886.8; DB 11; Length 1024;
Best local Similarity 99.8%; Pred. No. 76-255;
Matches 888; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAGATGCAATCTCTAAGAGAGCTTTGAAGAAATAAAGAAATGCGACGCTGGA 60
Db 108 AAGATGCAATCTCTAAGAGAGCTTTGAAGAAATAAAGAAATGCGACGCTGGA 167
Qy 61 GATGTTGCTAAAGCAATCACTAGCTGTTTATGTTAAAGCCAGACGATCTAT 120
Db 168 GATGTTGCTAAAGCAATCACTAGCTGTTTATGTTAAAGCCAGACGATCTAT 227
Qy 121 GAGGATTTGCACTTCTGTTGATATCTGTTGAGCCAGACGATGAGTCTCCAGAACCTTA 180
Db 228 GAGGATTTGCACTTCTGTTGATATCTGTTGAGCCAGACGATGAGTCTCCAGAACCTTA 287

Qy 181 GAAAAAGCATTCAAATTAATGATCAAAACCTGACGAGATGGGCTGAGAAAGTTTAC 240
Db 288 GAAAAAGCATTCAAATTAATGATCAAAACCTGACGAGATGGGCTGAGAAAGTTTAC 347
Qy 241 CTGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAAGATCAGTGTGATCTGAG 300
Db 348 CTGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAAGATCAGTGTGATCTGAG 407
Qy 301 CCAAGATTCATTAAGATAGCATTCCTGGTCTTGGGAGAGCATTTGGGATCTCCCGAA 360
Db 408 CCAAGATTCATTAAGATAGCATTCCTGGTCTTGGGAGAGCATTTGGGATCTCCCGAA 467
Qy 361 GGCATTACAGAGAGAGTTGTTGTTGATGACCTTTTCATGATGACAGAGGGCTCA 420
Db 468 GGCATTACAGAGAGAGTTGTTGTTGATGACCTTTTCATGATGACAGAGGGCTCA 527
Qy 421 GAACCAAGAGGAAAGATTGTTTATTAACCAACCTTACATCACTCAAGAGCGTG 480
Db 528 GAACCAAGAGGAAAGATTGTTTATTAACCAACCTTACATCACTCAAGAGCGTG 587
Qy 481 CAATACCGAAGCGAGGGGGGGGAGAGCTGCCAAGTGGGGGCTTGGCATCTCAT 540
Db 588 CAATACCGAAGCGAGGGGGGGGAGAGCTGCCAAGTGGGGGCTTGGCATCTCAT 647
Qy 541 CGATCCGTGGGCTCTCTCCATCTACATCTCAAGTCAAGATTTACAGATACAGAT 600
Db 648 CGATCCGTGGGCTCTCTCTCCATCTACATCTCAAGTCAAGATTTACAGATACAGAT 707
Qy 601 GGCCTGCCCAAAATTCACAGCCTGTATTACGGTGAAGATGCAAGATGATGTCAGA 660
Db 708 GGCCTGCCCAAAATTCACAGCCTGTATTACGGTGAAGATGCAAGATGATGTCAGA 767
Qy 661 ATGCTCTCATGGAGATCAAAATTTGATCAGTAAAGATGGGGGAAAGACTTACCA 720
Db 768 ATGCTCTCATGGAGATCAAAATTTGATCAGTAAAGATGGGGGAAAGACTTACCA 827
Qy 721 GATCTGATTCCTTCAACACTGTAGCAGATCACTGGAGCAAAATTTCCAGAACAGTT 780
Db 828 GATCTGATTCCTTCAACACTGTAGCAGATCACTGGAGCAAAATTTCCAGAACAGTT 887
Qy 781 GATCTGATTCCTTCAACACTGTAGCAGATCACTGGAGCAAAATTTCCAGAACAGTT 840
Db 888 GATCTGATTCCTTCAACACTGTAGCAGATCACTGGAGCAAAATTTCCAGAACAGTT 947
Qy 841 GGTGAGCCTTTATATGATGGAAGCATCTCACTTATTAAGATCTTG 890
Db 948 GGTGAGCCTTTATATGATGGAAGCATCTCACTTATTAAGATCTTG 997

RESULT 5
EX439467
LOCUS
DEFINITION Homo sapiens PLACENTA Homo sapiens cDNA clone CSDB010Y006
5-PRIME, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Li, W.-B., Gruber, C., Jeesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
<http://www.genoscope.cns.fr/>

cg1-bin/cluster.cgi?seq=CS0DE010DE3QPlc1cluster=4663.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DE010DE3QPl.

FEATURES

SOURCE

Location/Qualifiers
 1. 1050

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE010Y106"
 /issue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."
 BASE COUNT 307 a 233 c 271 g 239 t
 ORIGIN

Query Match 62.0%; Score 833.8; DB 13; Length 1050;
 Best Local Similarity 99.8%; Pred. No. 6.2e-239;
 Matches 835; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AAGATGGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGA 60
 214 AAGATGGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGA 273
 61 GATGTGCTAAGCAATCATCACTAGCTGTTATGTGTAAGCCAGCAAGATCTAT 120
 274 GATGTGCTAAGCAATCATCACTAGCTGTTATGTGTAAGCCAGCAAGATCTAT 333
 121 GAGGATGGGACCTCTGTTGATATCTGTGAGCCAGACGAGTGGCTCCAGAACCTA 180
 334 GAGGATGGGACCTCTGTTGATATCTGTGAGCCAGACGAGTGGCTCCAGAACCTA 393
 181 GAAAAAGCATCAATATATGATCCAAACCTGACGAAGATGGCTGAGAAAGTTTAC 240
 394 GAAAAAGCATCAATATATGATCCAAACCTGACGAAGATGGCTGAGAAAGTTTAC 453
 241 CTGAGCCAGTGAATATCCCACTGGGAGAGGGGAGAAATCAGCTGTGATGTGAG 300
 454 CTGAGCCAGTGAATATCCCACTGGGAGAGGGGAGAAATCAGCTGTGATGTGAG 513
 301 CCAAGAAATTCATAGATAGCATCTGCTGCTGSCACAGATGGGAGCTCTCCAGAA 360
 514 CCAAGAAATTCATAGATAGCATCTGCTGCTGSCACAGATGGGAGCTCTCCAGAA 573
 361 GGCATTACAGCAAGATCTGCTGCTGCTCTTTCATGAACTGACAGAAAGGCTCA 420
 574 GGCATTACAGCAAGATCTGCTGCTGCTCTTTCATGAACTGACAGAAAGGCTCA 633
 421 GAAGCAAGAGGAAATGTTGTTTATACCAACCTTACATCACTACTCAAGACGGTG 480
 634 GAAGCAAGAGGAAATGTTGTTTATACCAACCTTACATCACTACTCAAGACGGTG 693
 481 CAATACCAAGAGGAGGGGGGGTGAAGCTGCAAGTGGGGCTTGGCATCTTCACT 540
 694 CAATACCAAGAGGAGGGGGGGTGAAGCTGCAAGTGGGGCTTGGCATCTTCACT 753
 541 CGATCCGTGGCTCTCTTCATCTACAGTCTCTCAACAGATTCAGGAATACAGGAT 600
 754 CGATCCGTGGCTCTCTTCATCTACAGTCTCTCAACAGATTCAGGAATACAGGAT 813
 601 GAGGTGCCCAAAATTCACAGCTGTATTCAGGTGAAGATGCAAGATGATCAAGA 660
 814 GAGGTGCCCAAAATTCACAGCTGTATTCAGGTGAAGATGCAAGATGATCAAGA 873
 661 ATGGCTTCATAGGATCAAAATTCATCTACAGTGAAGATGGGGGCAAGACCTAACA 720
 874 ATGGCTTCATAGGATCAAAATTCATCTACAGTGAAGATGGGGGCAAGACCTAACA 933
 721 GATCTGATTCCTTCAACACTGTACAGATCACTGGGAGCAAAATATTCAGAAAGGTT 780

Db 934 GATCTGATTCCTTCAACACTGTACAGATCACTGGGAGCAAAATATTCAGAAAGGTT 993
 Qy 781 GATCTGATTCCTTCAACACTGTACAGATCACTGGGAGCAAAATATTCAGAAAGGTT 837
 Db 994 GATCTGATTCCTTCAACACTGTACAGATCACTGGGAGCAAAATATTCAGAAAGGTT 1050

RESULT 6
 LOCUS BX355940 1038 bp mRNA linear EST 05-MAY-2003
 DEFINITION BX355940 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 ACCESSION BX355940
 VERSION BX355940.1 GI:30384019
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to this cluster, see
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cg1-bin/cluster.cgi?seq=CS0D1006AB08QPlc1cluster=4663.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1006AB08QPl.

FEATURES

SOURCE

Location/Qualifiers
 1. 1038

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1006Y115"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoRV
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 304 a 237 c 262 g 233 t 2 others
 ORIGIN

Query Match 57.4%; Score 771.8; DB 13; Length 1038;
 Best Local Similarity 99.6%; Pred. No. 2.8e-220;
 Matches 784; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 AAGATGGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGA 60
 252 AAGATGGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGA 311
 61 GATGTGCTAAGCAATCATCACTAGCTGTTATGTGTAAGCCAGCAAGATCTAT 120
 312 GATGTGCTAAGCAATCATCACTAGCTGTTATGTGTAAGCCAGCAAGATCTAT 371
 121 GAGGATGGGACCTCTGTTGATATCTGTGAGCCAGACGAGTGGCTCCAGAACCTA 180
 372 GAGGATGGGACCTCTGTTGATATCTGTGAGCCAGACGAGTGGCTCCAGAACCTA 431
 181 GAAAAAGCATCAATATATGATCCAAACCTGACGAAGATGGGCTGAGAAAGTTTAC 240
 432 GAAAAAGCATCAATATATGATCCAAACCTGACGAAGATGGGCTGAGAAAGTTTAC 491
 241 CTGAGCCAGTGAATATCCCACTGGGAGAGGGGAGAAATCAGCTGTGATGTGAG 300

Db 492 CTGAGCCAGTAGAGATACCCCACTGGAGAGAGGAGAAATACAGCTGTGATGCTGGAG 551
 QY 301 CCAAGATTCATAGATAGACCACTCTGGCTTGGCAGACGATGGGACTCTCCACAGA 360
 Db 552 CCAAGATTCATAGATAGACCACTCTGGCTTGGCAGACGATGGGACTCTCCACAGA 611
 QY 361 GGCATTACAGCAGAGTTGTGTGTGATGACCTTTTCGATGAACCTGAGAGAGGGCTCA 420
 Db 612 GGCATTACAGCAGAGTTGTGTGTGATGACCTTTTCGATGAACCTGAGAGAGGGCTCA 671
 QY 421 GAGAGCAGAGGAGAGATTTGTGTGTATACCAACTTATACATCACTACTCAAGAGCGTG 480
 Db 672 GAGAGCAGAGGAGAGATTTGTGTGTATACCAACTTATACATCACTACTCAAGAGCGTG 731
 QY 481 CAATACCGAGCCAGAGGGGGGGTGGAGAGTGCACAGGTGGGGGCTTTGGCATCTCTCAT 540
 Db 732 CAATACCGAGCCAGAGGGGGGGTGGAGAGTGCACAGGTGGGGGCTTTGGCATCTCTCAT 791
 QY 541 CGATCCGTGGCTCTCTTCTCATCTACAGTCTCTCACAGATATTCAGAGATACAGAGAT 600
 Db 792 CGATCCGTGGCTCTCTTCTCATCTACAGTCTCTCACAGATATTCAGAGATACAGAGAT 851
 QY 601 GGCCTGCCCCAAATTCACACAGCTGTATTACGCTGGAAGATGCAGAAATGATTCAGAGA 660
 Db 852 GGCCTGCCCCAAATTCACACAGCTGTATTACGCTGGAAGATGCAGAAATGATTCAGAGA 911
 QY 661 ATGGCTTCTCATGGGATCAAAATTTGATTCAGCTTAAAGAT-GGGGGGCAAGAACCTACCC 719
 Db 912 ATGGCTTCTCATGGGATCAAAATTTGATTCAGCTTAAAGATGGGGGGCAAGAACCTACCC 971
 QY 720 AGATACGTATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAGACAGT 779
 Db 972 AGATACGTATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAGACAGT 1031
 QY 780 TGTACTG 786
 Db 1032 TGTACTG 1038

RESULT 7 866 bp mRNA linear EST 03-SEP-2002
 LOCUS BU146905
 DEFINITION AGENCOURT_7978204_lupski_dorsal_root_ganglion_Homo_sapiens_cDNA
 ACCESSION BU146905
 VERSION BU146905.1 GI:22660437
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 866)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.lnl.gov
 Plate: LLM13575 row: 1 column: 02
 High quality sequence stop: 731.
 Location/Qualifiers
 1..866
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6185233"
 /sex="male"

/tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="MDH108"
 /clone_lib="lupski_dorsal_root_ganglion"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
 NotI; Site 2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGCGCTCG-3' and
 5'-GACTAGTCTAGTCGAGCGGCGCCCT(15)-3', size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

BASE COUNT 257 a 182 c 233 g 191 t 3 others
 ORIGIN
 Query Match 55.1%; Score 740.6; DB 13; Length 866;
 Best Local Similarity 99.1%; Pred. No. 6.3e-211;
 Matches 765; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
 1 AAGATGGCATCTCTTAAGAGACTTT-TGAAGAAATAAAGAAATAGCCACTGTGG 59
 Db 9 AAGATGGCATCTCTTAAGAGACTTTTGAAGAAATAAAGAAATAGCCACTGTGG 68
 QY 60 AGATGTTGCTTAAGCAATCATCACTAGCTGTTTATGTTAAGTAAAGCCAGAACAGATCCTA 119
 Db 69 AGATGTTGCTTAAGCAATCATCACTAGCTGTTTATGTTAAGTAAAGCCAGAACAGATCCTA 128
 QY 120 TGAGCATTTGGCACTTCTGTGTGATGATGTTGAAGCCAGACTGAGTGGCTCCAGAACT 179
 Db 129 TGAGCATTTGGCACTTCTGTGTGATGATGTTGAAGCCAGACTGAGTGGCTCCAGAACT 188
 QY 180 AGAAAAAGCCATCCAAATTAATGATACAAAACCTGACAGAGATGGCTGAGAGAAATTTCA 239
 Db 189 AGAAAAAGCCATCCAAATTAATGATACAAAACCTGACAGAGATGGCTGAGAGAAATTTCA 248
 QY 240 CCTGAGCCAGATGAGATATACCCCACTGGAGAGAGAGAGAGAAATACAGCTGATGCTTGA 299
 Db 249 CCTGAGCCAGATGAGATATACCCCACTGGAGAGAGAGAGAGAAATACAGCTGATGCTTGA 308
 QY 300 GCCAAGATTTATAGATAGATAGCATCTGGGCTTTGGAGCAGACTTGGGACTCTCCAGA 359
 Db 309 GCCAAGATTTATAGATAGATAGCATCTGGGCTTTGGAGCAGACTTGGGACTCTCCAGA 368
 QY 360 AGGATTTACAGCAAAAGTTCTGTGTGATGATCTCTTTGATTAATCTGACAGAGAGGGCTC 419
 Db 369 AGGATTTACAGCAAAAGTTCTGTGTGATGATCTCTTTGATTAATCTGACAGAGAGGGCTC 428
 QY 420 AGAAGCAAGAGAGAGATGTTGTTTAAACCAACTTACATCACTACTCAAGAGAGCGT 479
 Db 429 AGAAGCAAGAGAGAGATGTTGTTTAAACCAACTTACATCACTACTCAAGAGAGCGT 488
 QY 480 GCAATACCGAAGCGAGGGGGGGTGAAGCTGCCAAGGTGGGGCTTTGGCATCTCTCAT 539
 Db 489 GCAATACCGAAGCGAGGGGGGGTGAAGCTGCCAAGGTGGGGCTTTGGCATCTCTCAT 548
 QY 540 TCGATCCGTGGCTCTCTTCTCATCTACAGTCTCTCAACAGATATTGAGAAATTCAGAGA 599
 Db 549 TCGATCCGTGGCTCTCTTCTCATCTACAGTCTCTCAACAGATATTGAGAAATTCAGAGA 608
 QY 600 TGGGGTGGCCCAAAATTCACACAGCTGTATTAAGGTTGAGAGATGAGAAATGATGTCAG 659
 Db 609 TGGGGTGGCCCAAAATTCACACAGCTGTATTAAGGTTGAGAGATGAGAAATGATGTCAG 668
 QY 660 AATGGCTTCTCATAGAGATCAAAATTTGATTCAGCTTAAGATGGGGCAAGACCTTACC 719
 Db 669 AATGGCTTCTCATAGAGATCAAAATTTGATTCAGCTTAAGATGGGGCAAGACCTTACC 728
 QY 720 AGATACGTATTCCTTCAACACT-GTACAGAGATCACTGGAGAGAAATATTC 770
 Db 729 AGATACGTATTCCTTCAACACTGTACAGAGATCACTGGAGAGAAATATTC 780

RESULT 8
LOCUS BG751497
DEFINITION 60273007F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4873770 5', mRNA sequence.
ACCESSION BG751497
VERSION BG751497.1 GI:14062150
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS 1 (bases 1 to 902)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-romail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLML at: http://image.llnl.gov
 Plate: LICM1752 row: 9 column: 19
 High quality sequence stop: 817.
FEATURES
 source
 1..902
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4873770"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH MGC 43"
 /note="Organ: eye; Vector: pOT87; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 246 a 200 c 251 g 204 t 1 others
ORIGIN
 Query Match 55.1%; Score 740; DB 10; Length 902;
 Best Local Similarity 95.7%; Pred. No. 9,6e-211;
 Matches 847; Conservative 0; Mismatches 26; Indels 12; Gaps 8;

259 CCCCACTGGAGAGAGGAGAAATCAAGCTGTGATGTCGAGCCAGAAATTCATPAAGATA 318
 2 CCCCACTGGAGAGAGGAGAAATCAAGCTGTGATGTCGAGCCAGAAATTCATPAAGATA 61
 319 GCCATCTCTGGGTCTTGGCAGAGCATGGGACTCTCTCAAGAGCATTACGCGAAGTT 378
 62 GCCATCTCTGGGTCTTGGCAGAGCATGGGACTCTCTCAAGAGCATTACGCGAAGTT 121
 379 CTGGTGTGACCTCTTTCGATGATGACGAGAGAGGCTCAGAAAGCAGAGGAGATT 438
 122 CTGGTGTGACCTCTTTCGATGATGACGAGAGAGGCTCAGAAAGCAGAGGAGATT 180
 439 GTTGTATTAAACCACTTACATCACTACTCAAGAGAGGATGCAATCCGAGCAGAGGG 498
 181 GTTGTATTAAACCACTTACATCACTACTCAAGAGAGGATGCAATCCGAGCAGAGGG 240
 499 GCGGTGAAGCTGCCAAGGTGGGGCTTGGCAGCTCTCATTCGATCGTGGCTCTTC 558
 241 GCGGTGAAGCTGCCAAGGTGGGGCTTGGCAGCTCTCATTCGATCGTGGCTCTTC 300
 559 TCATCTACAGTCTCAACAGGATATTCAGGAATACAGATGGGTCGCCAATATTCGA 618

RESULT 9
LOCUS BX355939/c
DEFINITION BX355939 Homo sapiens PIACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSOD1006115 3-PRIME, mRNA sequence.
ACCESSION BX355939
VERSION BX355939.1 GI:30382027
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS 1 (bases 1 to 996)
TITLE Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSOD1006A08NP1cluster=4663.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSOD1006A08NP1.

FEATURES
 source
 1..996
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

Qy 361 GGCAATTAAGAGCAAGATTCTGTGTGATGACCTCTTTGATGTAAGTCAGAGAGAGGCGCTCA 420
Db 496 GGCAATTAAGAGCAAGATTCTGTGTGATGACCTCTTTGATGTAAGTCAGAGAGAGGCGCTCA 555
Qy 421 GAAGCAAGAGAGAGATTGTTGTTTATTAACCACTTATCATCAACTACTCAAGAGAGCGTG 480
Db 556 GAAGCAAGAGAGAGATTGTTGTTTATTAACCACTTATCATCAACTACTCAAGAGAGCGTG 615
Qy 481 CAATACCGAAGCGAGGCGGCGGTGGAAGCTGCAAGGAGGCGGCTTTGGCATCTCTCATTT 540
Db 616 CAATACCGAAGCGAGGCGGCGGTGGAAGCTGCAAGGAGGCGGCTTTGGCATCTCTCATTT 675
Qy 541 CGATCGGTGGCTCTCTCTCATCTACAGCTCTCAACAGGTATTCAGAAATACAGAGAT 600
Db 676 CGATCGGTGGCTCTCTCTCATCTACAGCTCTCAACAGGTATTCAGAAATACAGAGAT 735
Qy 601 GGGGTGCCCCAAATTTCCACAGCCTGTATTACGGTGAAGATGCAAAATATGTCAAGA 660
Db 736 GGGGTGCCCCAAATTTCCACAGCCTGTATTACGGTGAAGATGCAAAATATGTCAAGA 795
Qy 661 ATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGCTAAAGAT-GGGGGCAAGACCTACCC 719
Db 796 ATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGGCAAGACCTACCC 855
Qy 720 AGATACTGATTCCTTCAACACT-GTAGCAGAGATCACTGGG 759
Db 856 AGATACTGATTCCTTCAACACTGTGTAGCAGAGATCACTGGG 896

RESULT 11
LOCUS CB958693 739 bp mRNA linear EST 29-APR-2003
DEFINITION IMAGE:30351744 5', mRNA sequence.
ACCESSION CB958693
VERSION CB958693.1 GI:30214809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 739)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovics
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: NDCM149 row: m column: 01
High quality sequence stop: 572.
Location/Qualifiers

FEATURES

SOURCE

1. 739
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30351744"
/lab_host="DH10B (TI phage-resistant)"
/clone_id="NIH_MGC_184"
/note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site 1:
SfiI (ggccatcggc); Site 2: SfiI (ggccggccggc);
Library is oligo-dr primed and directionally cloned. cDNA
parathyroid, adrenal, cortex and pineal gland. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTAATGGCC-3' and 3' adaptor sequence:
5'-ATTGAGAGCCGAGGCGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.38

kb (range 0.60-3.5 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH-MGC Library."

BASE COUNT

192 a 171 c 195 g 181 t

ORIGIN

Query Match 53.5%; Score 719.2; DB 14; Length 739;
Best Local Similarity 99.6%; Pred. No. 1.6e-204;
Matches 721; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 516 GGTGGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCA 575
Db 3 GGTGGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCA 62
Qy 576 CACAGTATTCAGGAATACCAAGATGGCGTCCCAAAATTCACAGCCTGTATTAGGT 635
Db 63 CACAGTATTCAGGAATACCAAGATGGCGTCCCAAAATTCACAGCCTGTATTAGGT 122
Qy 636 GGAAGATGCAAGAAATGATGTCAAGATGGCTTCATGGGGATCAAAATTTGTCATTCAGCT 695
Db 123 GGAAGATGCAAGAAATGATGTCAAGATGGCTTCATGGGGATCAAAATTTGTCATTCAGCT 182
Qy 696 AAGATGGGGCAAGACCTACCCAGATGATTCCTTCAACACTGTAGCAGAGATCAC 755
Db 183 AAGATGGGGCAAGACCTACCCAGATGATTCCTTCAACACTGTAGCAGAGATCAC 242
Qy 756 TGGGAGCAATATTCAGAAAGAGTTGATCTGTGATGATCTGACAGCTGGAGTGT 815
Db 243 TGGGAGCAATATTCAGAAAGAGTTGATCTGTGATGATCTGACAGCTGGAGTGT 302
Qy 816 TGGGAGAGGAGCCATGATGATGGCGGTGGAGCCTTATATCAAGGAGAGCTCTCACT 875
Db 303 TGGGAGAGGAGCCATGATGATGGCGGTGGAGCCTTATATCAAGGAGAGCTCTCACT 362
Qy 876 TATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTGCGTGGCTCTGATGACAG 935
Db 363 TATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTGCGTGGCTCTGATGACAG 422
Qy 936 AGAACAAGTGGAGTTGGTGTCTTCCAGATTTATTCATTCACAGGAATATTTCCAA 995
Db 423 AGAACAAGTGGAGTTGGTGTCTTCCAGATTTATTCATTCACAGGAATATTTCCAA 482
Qy 996 CTACAGTCTGTGATGAGTCTGACGAGAACTTTCCCACTGGCGTGCATTCAC 1055
Db 483 CTACAGTCTGTGATGAGTCTGACGAGAACTTTCCCACTGGCGTGCATTCAC 542
Qy 1056 TGGCAGTGAAGAGCCAGGCGCATCATGAGAGGATTATGAGCTGTGACAGCCCTCAA 1115
Db 543 TGGCAGTGAAGAGCCAGGCGCATCATGAGAGGATTATGAGCTGTGACAGCCCTCAA 602
Qy 1116 TATCACTCAGGTCTCTGAGCCATGAGAGAGGACAGACATCAACTTTTGATTCAGCTGG 1175
Db 603 TATCACTCAGGTCTCTGAGCCATGAGAGAGGACAGACATCAACTTTTGATTCAGCTGG 662
Qy 1176 AGTCCCTGAGCCGCTACTTGTGATCTTATACAGATTTCTTTCCATCTCCCA 1235
Db 663 AGTCCCTGAGCCGCTACTTGTGATCTTATACAGATTTCTTTCCATCTCCCA 722
Qy 1236 CGGA 1239
Db 723 CGGA 726

RESULT 12
LOCUS BX360507 1126 bp mRNA linear EST 05-MAY-2003
DEFINITION BX360507 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone GS001071A19 5-PRIME, mRNA sequence.
ACCESSION BX360507
VERSION BX360507.1 GI:30374439
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1126)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOD1071AA100P1cluster=4663.f. Contact :
Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSOD1071AA100P1.
FEATURES
source
1..1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1071YA19"
/issue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 311 a 249 c 298 g 246 t 22 others
ORIGIN
Query Match 52.3%; Score 702.4; DB 13; Length 1126;
Best Local Similarity 96.3%; Pred. No. 2.2e-199;
Matches 761; Conservative 9; Mismatches 15; Indels 5; Gaps 5;
QY 1 AAGATGGCATCTCTTAAGAGCATTTTGAAGAAATTAAGAAATGACCGCTGTGA 60
DB 282 AAGATGGCATCTCTTAAGAGCATTTTGAAGAAATTAAGAAATGACCGCTGTGA 341
QY 61 GATGTGTGAAGCATCATCAACCTAGCTGTTTATGTGAAGCCAGAACAGATCTAT 120
DB 342 GATGTGTGAAGCATCATCAACCTAGCTGTTTATGTGAAGCCAGAACAGATCTAT 401
QY 121 GAGCGATTGGCATCTCTGTGTGATACCTGTGAGCCAGACTGAGTCTCAAGACCTA 180
DB 402 GAGCGATTGGCATCTCTGTGTGATACCTGTGAGCCAGACTGAGTCTCAAGACCTA 461
QY 181 GAAAAAGCATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
DB 462 GAAAAAGCATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 521
QY 241 CTGAGCAGATGAGATACCCCACTGGAGAGAGAGAGAGAGATCACTGTGATGCTGAG 300
DB 522 CTGAGCAGATGAGATACCCCACTGGAGAGAGAGAGAGATCACTGTGATGCTGAG 581
QY 301 CCAAGATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 360
DB 582 CCAAGATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 641
QY 361 GGCATTACAGCAAGAGTTCTGTGTGATACCTCTTTTGAAGAACTCAGAGAGAGGCTCA 420
DB 642 GGCATTACAGCAAGAGTTCTGTGTGATACCTCTTTTGAAGAACTCAGAGAGAGGCTCA 701
QY 421 GAAAGCAGAGAGAGAGTTGTTTATTAACAACCTTACATCACTCAAGACGAGG 480
DB 702 GAAAGCAGAGAGAGAGTTGTTTATTAACAACCTTACATCACTCAAGACGAGG 761
QY 481 CAATACCGAAGCAGAGAGAGTTGTTTATTAACAACCTTACATCACTCAAGACGAGG 540
DB 762 CAATACCGAAGCAGAGAGAGTTGTTTATTAACAACCTTACATCACTCAAGACGAGG 821

QY 541 CGATCCGAGGCTCTCTTCCATCTACAGTCTGACAGAGTATTCAGAAATACAGAT 600
DB 822 CGATCCGAGGCTCTCTTCCATCTACAGTCTGACAGAGTATTCAGAAATACAGAT 881
QY 601 GCGGTGCCCAAAATTCACAGAGCTGTATTAACGCT-GGAAGATCAGAAATGATGTCAG 659
DB 882 GCGGTGCCCAAAATTCACAGAGCTGTATTAACGCTGTGGAAGATGACAAATGATGTCAG 941
QY 660 AATGCTTTCTTA-TGGATCAAAATTTGTCATTCAGCTTAAAGATGGGGCAAAACCTACC 718
DB 942 AATGCTTTCTTA-TGGATCAAAATTTGTCATTCAGCTTAAAGATGGGGCAAAACCTACC 1000
QY 719 CAGATCTGATTCCTTCAACAGCTGTAGAGATCACTGGAGCAAAATTCAGAACAG 778
DB 1001 CAGATCTGATTCCTTCAACAGCTGTAGAGATCACTGGAGCAAAATTCAGAACAG 1058
QY 779 TTGTACTGCT 788
DB 1059 TACTGKYAGT 1068
RESULT 13
BX355995 919 bp mRNA linear EST 02-MAY-2003
LOCUS BX355995
DEFINITION BX355995 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSOD1022Y16 5-PRIME, mRNA sequence.
ACCESSION BX355995
VERSION BX355995.1 GI:30339459
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 919)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOD1022DE08QP1cluster=4663.f. Contact :
Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSOD1022DE08QP1.
FEATURES
source
1..919
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1022Y16"
/issue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 273 a 206 c 226 g 208 t 6 others
ORIGIN
Query Match 51.7%; Score 694.2; DB 13; Length 919;
Best Local Similarity 98.3%; Pred. No. 5.8e-197;
Matches 707; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
QY 1 AAGATGGCATCTCTTAAGAGCATTTTGAAGAAATTAAGAAATGACCGCTGTGA 60
DB 198 AAGATGGCATCTCTTAAGAGCATTTTGAAGAAATTAAGAAATGACCGCTGTGA 257

QY 61 GATGTCCTAAGCAATCATCAACCTTACTGTTATAGTAAAGCCAGAACAGATCTTAT 120
 DB 258 GATGTCCTAAGCAATCATCAACCTTACTGTTATAGTAAAGCCAGAACAGATCTTAT 317
 QY 121 GAGGATTTGGCACTTCTGTTGATACCTGTTGAGCCAGACTGATGGCTCCAAAGACTTA 180
 DB 318 GAGGATTTGGCACTTCTGTTGATACCTGTTGAGCCAGACTGATGGCTCCAAAGACTTA 377
 QY 181 GAAAAAGCCATCCAAATTTATGACCAAAACCTGAGAGAAAGTGGGCTGGAGAAAGTTTAC 240
 DB 378 GAAAAAGCCATCCAAATTTATGACCAAAACCTGAGAGAAAGTGGGCTGGAGAAAGTTTAC 437
 QY 241 CTGAGGCAAGTGAAGATACCCCACTGGAGAGAGAGAGAAATCAGCTGTGATCTGGAG 300
 DB 438 CTGAGGCAAGTGAAGATACCCCACTGGAGAGAGAGAGAAATCAGCTGTGATCTGGAG 497
 QY 301 CCAAGAAATTCATAGATAGCCATCTCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGAA 360
 DB 498 CCAAGAAATTCATAGATAGCCATCTCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGAA 557
 QY 361 GGCATTCAGAGAAAGTTCTGGTGTGACCTCTTCCATGATGAACTGAGAGAGGGGCTCA 420
 DB 558 GGCATTCAGAGAAAGTTCTGGTGTGACCTCTTCCATGATGAACTGAGAGAGGGGCTCA 617
 QY 421 GAAAGCAAGAGGAGATTTGTTTATTAACCACTTACATCACTCAAGAGACGGTG 480
 DB 618 GAAAGCAAGAGGAGATTTGTTTATTAACCACTTACATCACTCAAGAGACGGTG 677
 QY 481 CAATACCGAAGCGAGGGGCGGTGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCAT 540
 DB 678 CAATACCGAAGCGAGGGGCGGTGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCAT 737
 QY 541 CGATCCGCGGCTCTCTTCCATCTACAGCTCTCAACAGATTTCCAGAAATACAGAGAT 600
 DB 738 CGATCCGCGGCTCTCTTCCATCTACAGCTCTCAACAGATTTCCAGAAATACAGAGAT 797
 QY 601 GGCCTGCGCCAAATTTCCACAGCTGTATTAACGTTGAGAGATGAGAAATGATGTCAGA 660
 DB 798 GGCCTGCGCCAAATTTCCACAGCTGTATTAACGTTGAGAGATGAGAAATGATGTCAGA 857
 QY 661 ATGGCTTCTCAT-GGATCAAAATTTGTCATTCAGCTAAAGATGGGGGCAAAAGACTTACC 718
 DB 858 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGCTAAABATGGGGGCAAAAGACTTACC 916

 RESULT 14
 CDS19131 813 bp mRNA linear EST 06-JUN-2003
 LOCUS AGENCCOURT 14372790 NIH MGC 181 Homo sapiens cDNA clone
 DEFINITION IMAGE:30400135 5', mRNA sequence.
 ACCESSION CDS19131
 VERSION CDS19131.1 GI:31450849
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 AUTHORS NIH-MGC
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgsdbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

FEATURES
 SOURCE
 1. 813
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30400135"
 /issue_type="White Matter"
 /dev_stage="Unknown"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistance)"
 /clone_idb="NIH MGC 181"
 /note="Vector: pCMV-Sport6.1; Site 1: NotI; Site 2: EcoRV
 (destroyed); Library is oligo-dT primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.42 kb. Library was constructed by
 (Invitrogen). Note: this is a NIH MGC Library."
 BASE COUNT 240 a 175 c 218 g 177 t 3 others
 ORIGIN
 Query Match 50.8%; Score 682.4; DB 14; Length 813;
 Best Local Similarity 98.2%; Pred. No. 2e-193;
 Matches 700; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
 QY 11 TCTCTAAGAGACTTTGAAGAAATAAAGAAATAGCAGCTGAGAGATTTGCTA 70
 DB 61 TCTAAGAGAGACTTTGAAGAAATAAAGAAATAGCAGCTGAGAGATTTGCTA 120
 QY 71 AAGCAATCATCAACCTAGCTGTTATGTTAAAGCCAGAACAGATCTTATGAGAGATTGG 130
 DB 121 AAGCAATCATCAACCTAGCTGTTATGTTAAAGCCAGAACAGATCTTATGAGAGATTGG 180
 QY 131 CACTTCTGTTGATCTGTTGAGCCAGACTGAGTGGCTCCAGAAATCTTGAAGAAAGCCA 190
 DB 181 CACTTCTGTTGATCTGTTGAGCCAGACTGAGTGGCTCCAGAAATCTTGAAGAAAGCCA 240
 QY 191 TCCAAATTAATGACCAAAACCTGAGAGAGTGGGCTGAGAAAGTCACTGGAGCCAG 250
 DB 241 TCCAAATTAATGACCAAAACCTGAGAGAGTGGGCTGAGAAAGTCACTGGAGCCAG 300
 QY 251 TGAAGATACCCCACTGGAGAGAGGAGAGAGATCAGCTGTGATGCTGGAGCCAGAAATTC 310
 DB 301 TGAAGATACCCCACTGGAGAGAGGAGAGAGATCAGCTGTGATGCTGGAGCCAGAAATTC 360
 QY 311 ATAAAGATAGCCATCTGAGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAATTAC 370
 DB 361 ATAAAGATAGCCATCTGAGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAATTAC 420
 QY 371 CAGAGATTCGAGGAGTCACTCTTTGATGATGATGAGAGAGAGGCTCCAGAAAGCAAGAG 430
 DB 421 CAGAGATTCGAGGAGTCACTCTTTGATGATGATGAGAGAGAGGCTCCAGAAAGCAAGAG 480
 QY 431 GGAAGATTTGTTTATTAACCACTTATCACTCACTCAAGAGCGTGCATATACCGAA 490
 DB 481 GGAAGATTTGTTTATTAACCACTTATCACTCACTCAAGAGCGTGCATATACCGAA 540
 QY 491 CGCAGGGGGCGGTGAGAACTGCGCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 550
 DB 541 CGCAGGGGGCGGTGAGAACTGCGCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 600
 QY 551 CTTCTCTCTCATCTAAGTCCACAGCTGATGATGAGAAATCCAGAGATGGCGTGGCCA 610
 DB 601 CTTCTCTCTCATCTAAGTCCACAGCTGATGATGAGAAATCCAGAGATGGCGTGGCCA 660
 QY 611 AAATTCAGAGCTGTATTAACGTTGAGAGATGAGAGATGATGTCAGAAATGGCTTCTC 670
 DB 661 AAATTCAGAGCTGTATTAACGTTGAGAGATGAGAGATGATGTCAGAAATGGCTTCTC 720
 QY 671 ATGGATCAAAATTTGTCATTCAGCTAAAGAT-GGGGCAAAAGACTTACCCAGA 722
 DB 721 ATGGATCAAAATTTGTCATTCAGCTAAAGATGGGGGCAAAAGACTTACCCAGA 773

RESULT 15
CD366462/c 708 bp mRNA linear EST 29-MAY-2003
LOCUS
DEFINITION
UI-H-F11-bjt-p-03-0-UI.s1 NCI CGAP_F11 Homo sapiens cDNA clone
CD366462
UI-H-F11-bjt-p-03-0-UI.3', mRNA sequence.
ACCESSION
CD366462.1 GI:31150552
VERSION
EST.
KEYWORDS
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 708)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/cgap.html>
Seq primer: M13 FORWARD
PolyA=Yes.
FEATURES
source
Location/Qualifiers
1..708
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-F11-bjt-p-03-0-UI"
/tissue_type="Aveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP F11"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI CGAP_F11 is a normalized cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pT73-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag
(dT)18 tail. The sequence tag for this library is
GGCCATGCCG. The tissue was provided by Dr. Gary W.
Hunninghake of the University of Iowa.
TAG LIB=UI-H-F11
TAG_TISSUE=Human Lung Aveolar Macrophage
TAG_SEQ=GGCCATGCCG"
BASE COUNT 158 a 192 c 149 g 209 t
ORIGIN
Query Match 50.7%; Score 681.8; DB 14; Length 708;
Best Local Similarity 99.7%; Pred. No. 2.8e-193;
Matches 663; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 588 TGGGCTTTGGCAGCAGATTGGAGCTCTCCAGAAAGGATTACGACGAAGTTCTGGTGG 529
QY 386 TGACCTCTTTTGATGAACCTGACAGAAAGGCTCTAGAAAGCAAGGGAATTGTGTTT 445
DB 528 TGACCTCTTTTGATGAACCTGACAGAAAGGCTCTAGAAAGCAAGGGAATTGTGTTT 469
QY 446 ATTAACAACCTTATCATGAATCACTCAAGAGCGGTGCAATACCGAAGCGAGGGCGGTGG 505
DB 468 ATTAACAACCTTATCATGAATCACTCAAGAGCGGTGCAATACCGAAGCGAGGGCGGTGG 409
QY 506 AAGCTGCCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTTCTCAATCT 565
DB 408 AAGCTGCCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTTCTCAATCT 349
QY 566 ACACTCTCTACACAGGATTTACAGAAATACCAAGATGCGTGGCCCAAAATTCCACAGCT 625
DB 348 ACAGTCTCTACACAGGATTTACAGAAATACCAAGATGCGTGGCCCAAAATTCCACAGCT 289
QY 626 GTATTACGGTGGAGATGCAAGAAATGATGTCAGAAATGATGCTTCATGGGATCAAAATTG 685
DB 288 GTATTACGGTGGAGATGCAAGAAATGATGTCAGAAATGATGCTTCATGGGATCAAAATTG 229
QY 686 TCATTACGCTTAAAGATGGGGCAAAAGCTTACCAGATGATGATTCCTTCAACACTGTAG 745
DB 228 TCATTACGCTTAAAGATGGGGCAAAAGCTTACCAGATGATGATTCCTTCAACACTGTAG 169
QY 746 CAGAGATCATCTGGAGCAAAATATCCAGAACAGTTGTACTGTACGTGACATCTGGACA 805
DB 168 CAGAGATCATCTGGAGCAAAATATCCAGAACAGTTGTACTGTACGTGACATCTGGACA 109
QY 806 GCTGGAGATGTTGGCAGGGTGCATGATGATGATGATGATGATGATGATGATGATGATGATG 865
DB 108 GCTGGAGATGTTGGCAGGGTGCATGATGATGATGATGATGATGATGATGATGATGATGATG 49
QY 866 CACTCTCACTTATTAAAGATCTTGG 890
DB 48 CACTCTCACTTATTAAAGATCTTGG 24

Search completed: December 22, 2003, 23:11:30
Job time : 2632.89 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:59:40 ; Search time 88.7694 Seconds
(without alignments)
6682.695 Million cell updates/sec

Title: US-09-745-763-35_COPY_171_1514
Perfect score: 1344
Sequence: 1 AAGAAATGCCTCTTAAGAG.....AAGAAATGCCTCTTAAGTCC 1344

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1340.8	99.8	1863	4	US-09-482-273-28 Sequence 28, Appl
2	772.4	57.5	1134	4	US-09-482-273-95 Sequence 95, Appl
3	48.2	3.6	7218	1	US-08-232-463-14 Sequence 14, Appl
4	37	2.8	2049	1	US-08-268-797-1 Sequence 1, Appl
5	37	2.8	2049	5	PCT-US95-08414-1 Sequence 1, Appl
6	37	2.8	2691	1	US-07-878-960-1 Sequence 1, Appl
7	35.2	2.6	2373	4	US-08-980-080-1 Sequence 1, Appl
8	35.2	2.6	3003	1	US-08-434-730-15 Sequence 15, Appl
9	34.6	2.6	4086	4	US-09-220-132-56 Sequence 56, Appl
10	34.6	2.6	4334	4	US-09-620-312D-88 Sequence 88, Appl
11	34.4	2.6	738	2	US-08-738-462-1 Sequence 1, Appl
12	34.4	2.6	738	5	PCT-US94-07587-1 Sequence 1, Appl
13	34	2.5	2106	4	US-09-252-991A-7477 Sequence 7477, Ap
14	34	2.5	2182	4	US-09-252-991A-7364 Sequence 7364, Ap
15	34	2.5	2187	4	US-09-252-991A-7167 Sequence 7167, Ap
16	33.4	2.5	2201	3	US-09-330-970-2 Sequence 2, Appl
17	33.4	2.5	2243	2	US-08-745-934-2 Sequence 4, Appl
18	33.2	2.5	2443	2	US-08-745-934-2 Sequence 2, Appl
19	33.2	2.5	9763	3	US-08-973-273-1 Sequence 1, Appl
20	33	2.5	1165	4	US-09-023-942A-28 Sequence 28, Appl
21	33	2.5	1443	4	US-09-328-352-2934 Sequence 2934, Ap
22	33	2.5	1491	4	US-09-328-352-1416 Sequence 1416, Ap
23	32.2	2.4	2272	4	US-09-313-394A-3342 Sequence 3342, Ap
24	32.2	2.4	2220	3	US-08-997-251-1 Sequence 1, Appl
25	32	2.4	4631	4	US-09-620-312D-164 Sequence 164, App
26	31.8	2.4	615	4	US-09-107-532A-1080 Sequence 1080, App
27	31.6	2.4	949	4	US-09-221-017B-594 Sequence 594, App

28	31.4	2.3	2242	3	US-09-400-742-1 Sequence 1, Appl
29	31.4	2.3	2242	3	US-08-618-651A-1 Sequence 1, Appl
30	31.4	2.3	2242	3	US-09-215-252-1 Sequence 1, Appl
31	31.2	2.3	1712	3	US-08-809-326A-8 Sequence 8, Appl
32	31.2	2.3	1712	4	US-09-689-914A-8 Sequence 8, Appl
33	31.2	2.3	1712	4	US-09-689-914A-8 Sequence 8, Appl
34	31.2	2.3	1712	4	US-09-689-914A-8 Sequence 8, Appl
35	31.2	2.3	1927	4	US-09-689-914A-13 Sequence 13, Appl
36	31.2	2.3	1927	4	US-09-689-914A-13 Sequence 13, Appl
37	31.2	2.3	1927	4	US-09-689-914A-13 Sequence 13, Appl
38	31.2	2.3	1954	3	US-08-809-326A-13 Sequence 13, Appl
39	31	2.3	280	3	US-08-783-853A-34 Sequence 34, Appl
40	31	2.3	280	3	US-09-344-050-34 Sequence 34, Appl
41	31	2.3	321	3	US-08-783-853A-43 Sequence 43, Appl
42	31	2.3	321	4	US-09-344-050-43 Sequence 43, Appl
43	31	2.3	1198	3	US-09-319-989-1 Sequence 1, Appl
44	31	2.3	162450	4	US-09-345-882-1 Sequence 1, Appl
45	30.8	2.3	1044	4	US-09-220-132-146 Sequence 146, App

ALIGNMENTS

```
RESULT 1
US-09-482-273-28
Sequence 28, Application US/09482273
Patent No. 6534631
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
CURRENT FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 28
LENGTH: 1863
TYPE: DNA
ORGANISM: Homo sapiens
US-09-482-273-28
Query Match 99.8%; Score 1340.8; DB 4; Length 1863;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 AAGATGCACTCTTAAGAGACTTTGAAGAAATAAAGAAATAGCAGCTGTGA 60
171 AAGATGCACTCTTAAGAGACTTTGAAGAAATAAAGAAATAGCAGCTGTGA 230
61 GATGTCCTTAAGCAATCATCAACTTACTGTTTATGTTAAAGCCAGAAATCTAT 120
231 GATGTCCTTAAGCAATCATCAACTTACTGTTTATGTTAAAGCCAGAAATCTAT 290
121 GAGGATGTCCTTAAGCAATCATCAACTTACTGTTTATGTTAAAGCCAGAAATCTAT 180
291 GAGGATGTCCTTAAGCAATCATCAACTTACTGTTTATGTTAAAGCCAGAAATCTAT 350
181 GAAAGGCAATCAATTAATGATCAAAACCTGACAGATGAGGCTGAGAAATTCAC 240
351 GAAAGGCAATCAATTAATGATCAAAACCTGACAGATGAGGCTGAGAAATTCAC 410
241 CTGAGGCAATGATGATATCCCTGAGAGAGGAGAGAGATGATGATGATGAGAG 300
411 CTGAGGCAATGATGATATCCCTGAGAGAGGAGAGAGATGATGATGATGAGAG 470
```

301 CCAAGATTCATTAAGATAGCCATCTGGGCTTTGGCAGCAGATGGGACTCTCCAGAA 360
471 CCAAGATTCATTAAGATAGCCATCTGGGCTTTGGCAGCAGATGGGACTCTCCAGAA 530
361 GGCATTACAGCAGAAAGTTCTGGTGTGACCTCTTTGATGAACTGCAAGAAAGGCTTCA 420
531 GGCATTACAGCAGAAAGTTCTGGTGTGACCTCTTTGATGAACTGCAAGAAAGGCTTCA 590
421 GAAGCAAGAGGAAATGTTGTTATATACCAACCTTACATCACTACTCAAGAGCGGTG 480
591 GAAGCAAGAGGAAATGTTGTTATATACCAACCTTACATCACTACTCAAGAGCGGTG 650
481 CAATCCCAAGCAGAGGCGGCTGGAGCTGCCAAGTGGGCGCTTGGCATCTCTCAT 540
651 CAATCCCAAGCAGAGGCGGCTGGAGCTGCCAAGTGGGCGCTTGGCATCTCTCAT 710
541 CGATCCGTCGCTCTCTTCCATCTCACTCACTCACTCACTCACTCACTCACTCACTCA 600
711 CGATCCGTCGCTCTCTTCCATCTCACTCACTCACTCACTCACTCACTCACTCACTCA 770
601 GCGCGGCCCAAAATTCACAGCCCTGTATACGGTGAAGATGCAAGATGATGTCAGA 660
771 GCGCGGCCCAAGATTCACAGCCCTGTATACGGTGAAGATGCAAGATGATGTCAGA 830
661 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGTAAAGATGGGCGCAAGACTTACCCA 720
831 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGTAAAGATGGGCGCAAGACTTACCCA 890
721 GATATGATTCCTTCAACAACCTGTAGCAGAGATCACTGGGAGCAAAATTCAGAACAGGTT 780
891 GATATGATTCCTTCAACAACCTGTAGCAGAGATCACTGGGAGCAAAATTCAGAACAGGTT 950
781 GTACTGTCAGTGAACATCTGAGCAGCTGGAGTGTGGGAGGCGCATGGATGATGATG 840
951 GTACTGTCAGTGAACATCTGAGCAGCTGGAGTGTGGGAGGCGCATGGATGATGATG 1010
841 GGTGAGGCTTTATATCATGGGAGCACTCTCATCTTAAAGATCTTGGGCTGCTCCA 900
1011 GGTGAGGCTTTATATCATGGGAGCACTCTCATCTTAAAGATCTTGGGCTGCTCCA 1070
901 AAGAGACTCTGCGGCTGCTGCTGCACTGCAAGAAACAAGGTGAGTGTGCTTC 960
1071 AAGAGACTCTGCGGCTGCTGCTGCACTGCAAGAAACAAGGTGAGTGTGCTTC 1130
961 CAGTATTCATGATTCACAAAGTAAATTTTCCACTACAGTCTGATGATGATGATGATG 1020
1131 CAGTATTCATGATTCACAAAGTAAATTTTCCACTACAGTCTGATGATGATGATGATG 1190
1021 GCAGGAACCTTCTTACCACTGGGCTGCAATTCATGGCAGTGAAGAGCCAGGCGCATC 1080
1191 GCAGGAACCTTCTTACCACTGGGCTGCAATTCATGGCAGTGAAGAGCCAGGCGCATC 1250
1081 ATGAGAGAGTTATGAGCTGCTGAGCCCTCAATATCACTCAGTCTTGAAGCCATGGA 1140
1251 ATGAGAGAGTTATGAGCTGCTGAGCCCTCAATATCACTCAGTCTTGAAGCCATGGA 1310
1141 GAAGGAGAGCATCAATCTTGGATCCAAGTGAAGTCCGAGGCGCATCACTTAT 1200
1311 GAAGGAGAGCATCAATCTTGGATCCAAGTGAAGTCCGAGGCGCATCACTTAT 1370
1201 GACTTATACAGTATTTCTTCTTCATCACTCCACGAGACACCATACATGATGAT 1260
1371 GACTTATACAGTATTTCTTCTTCATCACTCCACGAGACACCATACATGATGAT 1430
1261 CCAAGAGAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1320
1431 CCAAGAGAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1490
1321 ATGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1344
1491 ATGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1514

RESULT 2
us-09-482-273-95
; Sequence 95: Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482, 273
; EARLIER FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092, 921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092, 922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092, 956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
us-09-482-273-95
Query Match 57.5%; Score 772.4; DB 4; Length 1134;
Best Local Similarity 99.7%; Pred. No. 6e-234;
Matches 784; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
559 TCCATCAGAGTCTCTCAACAGGATTCAGAAATACAGAGATGGCTGCCAAATTCGA 618
1 TCCATCAGAGTCTCTCAACAGGATTCAGAAATACAGAGATGGCTGCCAAATTCGA 60
619 ACAGCTGTATTCAGTGAAGATGCAAGATGATGTCAGAAATGCTTCTCATGGATC 678
61 ACAGCTGTATTCAGTGAAGATGCAAGATGATGTCAGAAATGCTTCTCATGGATC 120
679 AAAATTCATTCAGTGAAGATGCAAGATGATGTCAGAAATGCTTCTCATGGATC 738
121 AAAATTCATTCAGTGAAGATGCAAGATGATGTCAGAAATGCTTCTCATGGATC 180
739 ACTGTACAGAGATCACTGGAGCAATATTCAGAAACAGGTTGATCTGTCAGTGAAT 798
181 ACTGTACAGAGATCACTGGAGCAATATTCAGAAACAGGTTGATCTGTCAGTGAAT 240
799 CTGACAGCTGGATGTTGGGCAAGGTCGATGATGATGATGATGATGATGATGATGAT 858
241 CTGACAGCTGGATGTTGGGCAAGGTCGATGATGATGATGATGATGATGATGATGAT 300
859 TGGAGACATCTCTCACTTATTAAGATCTTGGGCTGCTGCTCAAGAGAGCTTGGGCTG 918
301 TGGAGACATCTCTCACTTATTAAGATCTTGGGCTGCTGCTCAAGAGAGCTTGGGCTG 360
919 GTGCTGCACTGCAAGAAACAAGTGAAGTGTGCTTCCAGTATTCAGTGAAT 978
361 GTGCTGCACTGCAAGAAACAAGTGAAGTGTGCTTCCAGTATTCAGTGAAT 420
979 AAGTAAATATTTCCAACTACAGTCTGATGATGATGATGATGATGATGATGATGAT 1038
421 AAGTAAATATTTCCAACTACAGTCTGATGATGATGATGATGATGATGATGATGAT 480
1039 ACTGGGCTGCAATCACTGCAAGTGAAGAAAGGCGGCGCATTCATGAGAGGTTATGAG 1098
481 ACTGGGCTGCAATCACTGCAAGTGAAGAAAGGCGGCGCATTCATGAGAGGTTATGAG 539
1099 CTGCTGAGCCCTCAATATCACTCAGGCTCTGAGCCATGAGAGAGGAGCAGATCAAC 1158
540 CTGCTGAGCCCTCAATATCACTCAGGCTCTGAGCCATGAGAGAGGAGCAGATCAAC 599
1159 TTTTGGATCCAAAGTGAAGTGTGCTGAGCCAGTCTTATGATGATGATGATGATGATG 1218
600 TTTTGGATCCAAAGTGAAGTGTGCTGAGCCAGTCTTATGATGATGATGATGATGATG 659

QY	1125	GGGTCTGAGCCATGTGAGAGAGG	CAGACATCACTTTTGGATCCAGCGTGAGATGCTCGG	1184
Db	1900	GGTCAGACAATTGGTGTATGACATGTGACCAAGCCATTGTGGGCATGATGTCAGGCTCGG	1841	

QY 1185 AGCCAGTCTACTGTGATGACTTATACAGATTTCTTCCATCACTCCCGAGAGACAC 1244
DB 1840 CAACAGGCTCTGTTGTGACACACACATGTTTTTCAAGTGACTTCCACTTGTAC 1781
QY 1245 CATGACTGTGATGATCCAAAGCAGATGATGTTGCTGCTGTTGGGCTGTTGTTTC 1304
DB 1780 CTTGGAGACCTTATGCGCACGAGGCCCGATGCTCGCTAACACAGATTCATCAGC 1721
QY 1305 TTATGTTGT 1313
DB 1720 CAATGTGT 1712

RESULT 5

PCT-US95-08414-1/C
Sequence 1, Application PC/TUS9508414
GENERAL INFORMATION:
APPLICANT: Purchio, Anthony F.
APPLICANT: Lebaron, Richard
TITLE OF INVENTION: Factor to Grow Tissue Ex Vivo
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08414
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/268,797
FILING DATE: July 1, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: TISSUE.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2049 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-08414-1.

Query Match 2.8%; Score 37; DB 5; Length 2049;
Best Local Similarity 49.7%; Pred. No. 0.29;
Matches 94; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1125 GGTCTGAGCCATGAGAGGAGACATCACTTTTGGATCCAGCTGGAGTGCCTGG 1184
DB 1900 GCTGAGAACATGTTGATGATGATGACACCGCATTTGTGCTCATGATGTACAGCTCGG 1841
QY 1185 AGCCAGTCTACTGTGATGACTTATACAGATTTCTTCCATCACTCCCGAGAGACAC 1244
DB 1840 CAACAGGCTCTGTTGTGACACACACATGTTTTTCAAGTGACTTCCACTTGTAC 1781
QY 1245 CATGACTGTGATGATCCAAAGCAGATGATGTTGCTGCTGTTGGGCTGTTGTTTC 1304

DB 1780 CTTGGAGACCTTATGCGCACGAGGCCCGATGCTCGCTAACACAGATTCATCAGC 1721
QY 1305 TTATGTTGT 1313
DB 1720 CAATGTGT 1712

RESULT 6

US-07-878-960-1/C
Sequence 1, Application US/07878960
Patent No. 5444164
GENERAL INFORMATION:
APPLICANT: Purchio, Anthony F.
APPLICANT: Skonier, John
APPLICANT: Neuberger, Michael G.
TITLE OF INVENTION: TGF-BETA INDUCED GENE AND PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company
STREET: 3005 First Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/878,960
FILING DATE: 05-MAY-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/833,835
FILING DATE: 05-FEB-1992
NAME: 530
ATTORNEY/AGENT INFORMATION:
NAME: Sorrentino, Joseph M.
REGISTRATION NUMBER: 32,598
REFERENCE/DOCKET NUMBER: ON0092-
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206/727-3601
TELEFAX: 206/728-4800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2691 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: LUNG
CELL TYPE: ADENOCARCINOMA
CELL LINE: A549
US-07-878-960-1

Query Match 2.8%; Score 37; DB 1; Length 2691;
Best Local Similarity 49.7%; Pred. No. 0.34;
Matches 94; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1125 GGTCTGAGCCATGAGAGGAGACATCACTTTTGGATCCAGCTGGAGTGCCTGG 1184
DB 1947 GCTCAGAACATTTGATGATGATGACACCGCATTTGTGCTCATGATGTACAGCTCGG 1888
QY 1185 AGCCAGTCTACTGTGATGACTTATACAGATTTCTTCCATCACTCCCGAGAGACAC 1244
DB 1887 CAACAGGCTCTGTTGTGACACACACATGTTTTTCAAGTGACTTCCACTTGTAC 1828
QY 1245 CATGACTGTGATGATCCAAAGCAGATGATGTTGCTGCTGTTGGGCTGTTGTTTC 1304

Db 1827 CTGGAGAGACTTACGCCGACCAAGGCCCGATGCTCGGCTAACAGAGATTTCATCAC 1768
QY 1305 TTATGTGT 1313
Db 1767 CATGTGT 1759

RESULT 7

US-08-980-080-1
Sequence 1, Application US/08980080
Patent No. 6312941
GENERAL INFORMATION:
APPLICANT: CARTER-SU, CHRISTIN
APPLICANT: RUI, LIANG-YOU
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
TITLE OF INVENTION: SIGNALING PATHWAY AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,080
FILING DATE: 26-NOV-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: UM-03036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2373 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-980-080-1

Query Match 2.6%; Score 35.2; DB 4; Length 2373;
Best Local Similarity 44.4%; Pred. No. 1.2; Indels 0; Gaps 0;
Matches 142; Conservative 0; Mismatches 178;

QY 771 AGACAGGTTGATGTCAGTGCATCTGACAGCTGGAGTGTGGGACGGTGCAT 830
Db 729 AGCAGAGTGTATACAGAGAGAGCTGCTGATTTATGGGGGTGAAGGCTGCC 788
QY 831 GGATGATGGCGGTGAGCTTTATATCATGGGAAGCACTTCACTTATTAAGATCTTGG 890
Db 789 TGACCCAGCAGAGATGCTGAGGAGGCGGAGCTGGCTGACCTCGGAGGAGGAGG 848
QY 891 GCTGCGTCAAGAGAGACTGCGGCTGCTGCTGCACTGCAAGAAAGAGTGAAGT 950
Db 849 GCACCTCAGTGGCAGAAATGTCACTGCTCGGAGTGAAGAGAGAGAGAGG 908
QY 951 TGTGCTTCCAGTATATATCACTTACACAGGTAAATTTTCCAACTACAGTCTGAT 1010
Db 909 AAGTGGCTTGAAGTTCTTTGACACCCAGGCAATCCCGGCGCTTTAGCATTTCCCTG 968
QY 1011 GGAAGTCAAGCAGAACTTTCTTACCACTGGGCTGCAATTCAGTGCAGTGAAGGC 1070
Db 969 TTCTACTATTACTGATGTCCGACAGCCAGCCCTGAGATGCTGACAGGAGAAACAC 1028

QY 1071 CAGGGCCATCATGAGAGAG 1090
Db 1029 GTTTGTGTTAAGGTAGAG 1048

RESULT 8

US-08-434-730-15
Sequence 15, Application US/08434730
Patent No. 5637463
GENERAL INFORMATION:
APPLICANT: Dalton, Stephen
APPLICANT: Kochan, Jarema P
APPLICANT: Osborne, Mark A
TITLE OF INVENTION: METHOD TO DETECT PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTIONS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,730
FILING DATE: 04-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Semionow, Raina
REGISTRATION NUMBER: 39022
REFERENCE/DOCKET NUMBER: 9069
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)235-4391
TELEFAX: (201)235-2363
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3003 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: YES
ANTI-SENSE: NO
US-08-434-730-15

Query Match 2.6%; Score 35.2; DB 1; Length 3003;
Best Local Similarity 44.4%; Pred. No. 1.3;
Matches 142; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 771 AGACAGGTTGATGTCAGTGCATCTGACAGCTGGAGTGTGGGACGGTGCAT 830
Db 1072 AGCAGAGTGTATACAGAGAGAGCTGCTGATTTATGGGGGTGAAGGCTGCC 1131
QY 831 GGATGATGGCGGTGAGCTTTATATCATGGGAAGCACTTCACTTATTAAGATCTTGG 890
Db 1132 TGACCCAGCAGAGATGCTGAGGAGGCGGAGCTGGCTGACCTCGGAGGAGG 1191
QY 891 GCTGCGTCAAGAGAGACTGCGGCTGCTGCTGCACTGCAAGAAAGAGTGAAGT 950
Db 1192 GCACCTCAGTGGCAGAAATGTCACTGCTCGGAGTGAAGAGAGAGAGAGG 1251
QY 951 TGTGCTTCCAGTATATATCACTTACACAGGTAAATTTTCCAACTACAGTCTGAT 1010
Db 1252 AAGTGGCTTGAAGTTCTTTGACACCCAGGCAATCCCGGCGCTTTAGCATTTCCCTG 1311
QY 1011 GGAAGTCAAGCAGAACTTTCTTACCACTGGGCTGCAATTCAGTGCAGTGAAGGC 1070
Db 1312 TTCTACTATTACTGATGTCCGACAGCCAGCCCTGGAATGCTGACAGGAGAAACAC 1371

Query Match 2.6%; Score 34.4; DB 2; Length 738;
Best Local Similarity 48.4%; Pred. No. 1;
Matches 124; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 624 CTGATTAACGTTGAAGATGAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAAT 683
DB 261 CTGATTAAGAGATGCTCACTGAAACATCCAGTAACAACAAGAGAAATAT 202
QY 684 TGTCACTACGTTAAAGATGGGGGCAAGACCTTACCCAGAT-ACGTATCTTCAACACTG 742
DB 201 CCCAGCAGAGTAAAGTTTTCAGGGCAAAATGATGCAAGAGTAACCTGACAGACATCCG 142
QY 743 TAGCAGATCACTGGAGCAAAATATCCAGACAGTTGTACTGTGATGAGATCTGG 802
DB 141 AGGAAGATGAAGTGAAGAACTTTCTGGGCTGAGTCTGTGGTAAGTTAACTAGC 82
QY 803 ACAGCTGGATGTTGGGCAAGGTGCCATGATGATGCGGTGGAGCCTTTATATCATGG 862
DB 81 ATATATTGCTGTGGTGCATGAGACAGAGAAAGGTGCATTAACTGTGTGTAAGA 22

QY 863 AAGCACTCTCACTTAT 878
DB 21 ACAAACTCTCAATTCT 6

RESULT 12
PCT-US94-07587-1/c
Sequence 1, Application PC/TUS9407587
GENERAL INFORMATION:
APPLICANT: Schering Corp.
TITLE OF INVENTION: PURIFIED MAMMALIAN NK ANTIGENS AND
TITLE OF INVENTION: RELATED REAGENTS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: One Giralda Farms
CITY: Madison
STATE: New Jersey
COUNTRY: USA
ZIP: 07940
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07587
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G.
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: DX0397K
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-822-7255
TELEFAX: 201-822-7039
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURES:
NAME/KEY: CDS
LOCATION: 61..738
PCT-US94-07587-1

Query Match 2.6%; Score 34.4; DB 5; Length 738;
Best Local Similarity 48.4%; Pred. No. 1;
Matches 124; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 624 CTGATTAACGTTGAAGATGAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAAT 683

DB 261 CTGATTAAGAGATGCTCACTGAAACATCCAGTAACAACAAGAGAAATAT 202
QY 684 TGTCACTACGTTAAAGATGGGGGCAAGACCTTACCCAGAT-ACGTATCTTCAACACTG 742
DB 201 CCCAGCAGAGTAAAGTTTTCAGGGCAAAATGATGCAAGAGTAACCTGACAGACATCCG 142
QY 743 TAGCAGATCACTGGAGCAAAATATCCAGACAGTTGTACTGTGATGAGATCTGG 802
DB 141 AGGAAGATGAAGTGAAGAACTTTCTGGGCTGAGTCTGTGGTAAGTTAACTAGC 82
QY 803 ACAGCTGGATGTTGGGCAAGGTGCCATGATGATGCGGTGGAGCCTTTATATCATGG 862
DB 81 ATATATTGCTGTGGTGCATGAGACAGAGAAAGGTGCATTAACTGTGTGTAAGA 22

QY 863 AAGCACTCTCACTTAT 878
DB 21 ACAAACTCTCAATTCT 6

RESULT 13
US-09-252-991A-7477
Sequence 7477, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
SEQ ID NO 7477
NUMBER OF SEQ ID NOS: 33142
LENGTH: 2106
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7477

Query Match 2.5%; Score 34; DB 4; Length 2106;
Best Local Similarity 54.9%; Pred. No. 2.6;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 970 CAGTTACACAAGTAAATATTTCACACTACAGTCTGTGATGAGTGTGACGACGAACC 1029
DB 1181 CAGGAAGTGAAGACGTGATGTCTTCTCGCTGTGGTGAACCCGACGACGAAC 1240
QY 1030 TTCTTACCACTGGGCTGCATTCACTGGCAGTGAAGGACGAGGCTATCATGAGAG 1089
DB 1241 GCCTTCTCAGGGTGAATCAAGTCCGCCGGGAAATCGTTGGGACCTTGAGAG 1300

QY 1090 GT 1091
DB 1301 CT 1302

RESULT 14
US-09-252-991A-7334
Sequence 7334, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190

```

: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 7334
: LENGTH: 2112
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7334

```

Query Match	2.5%	Score 34	DB 4	Length 2112
Best Local Similarity	54.9%	Pred. No. 2.6		
Matches 67	Conservative 0	Mismatches 55	Indels 0	Gaps 0

QY	970	CAGTTACCAAGTAAATATTTCCAACTACAGCTGCTGTGATGAGTCTGACGACAGAAAC	1029
Db	1243	CAGGAAGTGAAGGAGCACTGATGTCTCTACTTCCGCTCTGCTGTGTAACCCCGAGACACACAC	1302
QY	1030	TTCTTACCCACTGGGCTGCATTTCACTGGCAGTGAAGAGCCACGGGCATCATGGAGGAG	1089
Db	1303	GCCTTCTCAGAGGTGATCAACGTGCGCGCGGGAATCGGTTGGGCAACCTTGGAGAG	1362
QY	1090	GT	1091
Db	1363	CT	1364

RESULT 15
 US-09-252-991A-7167/c
 : Sequence 7167, Application US/09252991A
 : Patent No. 6551795
 : GENERAL INFORMATION:
 : APPLICANT: Marc J. Rubenfield et al.
 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 : TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 : FILE REFERENCE: 107196.136
 : CURRENT APPLICATION NUMBER: US/09/252,991A
 : CURRENT FILING DATE: 1999-02-18
 : PRIOR APPLICATION NUMBER: US 60/074,788
 : PRIOR FILING DATE: 1998-02-18
 : PRIOR APPLICATION NUMBER: US 60/094,190
 : PRIOR FILING DATE: 1998-07-27
 : NUMBER OF SEQ ID NOS: 33142
 : SEQ ID NO 7167
 : LENGTH: 2187
 : TYPE: DNA
 : ORGANISM: Pseudomonas aeruginosa
 : US-09-252-991A-7167

Query Match	2.5%	Score 34	DB 4	Length 2187
Best Local Similarity	54.9%	Pred. No. 2.7		
Matches 67	Conservative 0	Mismatches 55	Indels 0	Gaps 0

QY	970	AGTTACACAAGGTAAATATTTCCAACTACAGCTCGTGATGAGATCTGACGCAAGAAC	1029
Db	975	CAGGAAGTGAAGGACCTGATGTCTTACTTCGCGCTGTGTGAACCCGACGACGACAAAC	916
QY	1030	TTCTTACCCACTGGGCTGCATTTCACTGGCACTGAAGAAGCGCAGGGCCATCATGGAGAG	1089
Db	915	GCCCTTCCTCAAGAGGTATCAACGTCGCCGCGCGGAAATCGGTCTGGCCGACCTTGAGAAAG	856
QY	1090	GT	1091
Db	855	CT	854

Search completed: December 22, 2003, 23:17:21
Job time : 90.7694 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 14:49:05 ; Search time 420.487 Seconds
(without alignments)
10650.710 Million cell updates/sec

Title: US-09-745-763-35_COPY_171_1514
Perfect score: 1344
Sequence: 1 AAGATGTCATCTTAAGAG.....AAGATGTCCTAGGTCC 1344

Scoring table: IDENTITY NUC
Gapop 10.0, Gapept 1.0

Searched: 2211978 seqs, 1666101734 residues
Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 - 5: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
 - 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 - 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
 - 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
 - 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 - 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	1851	9	US-09-745-763-35
2	1340.8	99.8	1863	11	US-09-984-271-28
3	1060.8	78.9	1778	10	US-09-917-800A-505
4	772.4	57.5	1134	11	US-09-984-271-95
5	261.8	19.5	357	10	US-09-833-381-1929
6	215.8	16.1	427	10	US-09-833-381-1930
7	91	6.8	134	10	US-09-783-590-12141
8	60.2	4.5	65	13	US-09-908-975-25931
9	42	3.1	2200	15	US-09-908-975-8960
10	41	3.1	65	13	US-09-908-975-1202
11	39.4	2.9	484	11	US-09-918-995-30889
12	39.4	2.9	1363	15	US-10-232-484-5
13	39.4	2.9	2319	14	US-10-044-090-579
14	39.4	2.9	3322	13	US-09-814-353-20277
15	39.4	2.9			Sequence 20277, A

C 16	38.6	2.9	3997	15	US-10-198-846-12639	Sequence 12639, A
C 17	38	2.8	457	13	US-10-027-632-310889	Sequence 310889,
C 18	38	2.8	457	13	US-10-027-632-310889	Sequence 310889,
C 19	37.8	2.8	999	15	US-10-184-644-434	Sequence 434, App
C 20	37.8	2.8	999	15	US-10-184-644-434	Sequence 434, App
C 21	37	2.8	502	15	US-10-066-543-2028	Sequence 2028, App
C 22	37	2.8	548	15	US-10-066-543-2222	Sequence 2222, App
C 23	37	2.8	568	10	US-09-998-598-2545	Sequence 2545, Ap
C 24	37	2.8	2691	9	US-09-922-217-121	Sequence 121, App
C 25	37	2.8	2691	10	US-09-833-263-121	Sequence 121, App
C 26	37	2.8	2691	10	US-09-880-107-2399	Sequence 2399, Ap
C 27	37	2.8	2691	13	US-10-301-822-200	Sequence 200, App
C 28	37	2.8	2691	13	US-09-873-367C-1009	Sequence 1009, Ap
C 29	37	2.8	2691	14	US-10-025-524-1708	Sequence 121, App
C 30	37	2.8	2691	15	US-10-044-090-647	Sequence 1708, Ap
C 31	37	2.8	2782	14	US-10-084-817-28	Sequence 647, App
C 32	37	2.8	2782	15	US-10-208-408-16	Sequence 16, App1
C 33	37	2.8	2782	15	US-10-084-817-28	Sequence 28, App1
C 34	37	2.8	3313	13	US-09-814-353-19947	Sequence 19947, A
C 35	37	2.8	3320	15	US-10-198-846-11549	Sequence 11549, A
C 36	36.4	2.7	5085	15	US-10-198-846-9854	Sequence 9854, Ap
C 37	35.4	2.6	2574	13	US-10-311-626-13	Sequence 13, App1
C 38	35.2	2.6	1115	15	US-10-184-644-440	Sequence 440, App
C 39	35.2	2.6	1115	15	US-10-184-634-440	Sequence 440, App
C 40	34.6	2.6	416	11	US-09-918-995-16569	Sequence 16569, A
C 41	34.6	2.6	1237	13	US-10-027-632-199318	Sequence 199318,
C 42	34.6	2.6	1237	14	US-10-027-632-199318	Sequence 199318,
C 43	34.6	2.6	4182	15	US-10-060-036-66	Sequence 66, App1
C 44	34.6	2.6	4182	15	US-10-060-036-137	Sequence 137, App
C 45	34.6	2.6	4334	13	US-10-117-722-88	Sequence 88, App1

ALIGNMENTS

RESULT 1
US-09-745-763-35
Sequence 35, Application US/09745763
Patent No. US20020065394A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallee, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Werberg, David
Trecay, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 1851 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-745-763-35

Query Match 100.0%; Score 1344; DB 9; Length 1851;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAGATGGCATCTCTAAGAGACCTTTTGAAGAAATGAAGAAATGCGGCTGTGGA 60
DB 171 AAGATGGCATCTCTAAGAGACCTTTTGAAGAAATGAAGAAATGCGGCTGTGGA 230
QY 61 GATGTTGCTAAAGCATCATCACTAGCTGTTTATGGTAAAGCCGAGAACAGATCTAT 120
DB 231 GATGTTGCTAAAGCATCATCACTAGCTGTTTATGGTAAAGCCGAGAACAGATCTAT 290
QY 121 GAGCGATTGGACCTTCTGTTGATATCTGTGACCCAGACTGAGTGGCTCCAGAACCTA 180
DB 291 GAGCGATTGGACCTTCTGTTGATATCTGTGACCCAGACTGAGTGGCTCCAGAACCTA 350
QY 181 GAAAAAGCATCCAAATTTATGACAAAACCTGACGACAAATGGGCTGAGAAATTCAC 240
DB 351 GAAAAAGCATCCAAATTTATGACAAAACCTGACGACAAATGGGCTGAGAAATTCAC 410
QY 241 CTGGAGCCAGTGAATACTCCCACTGGAGAGAGAGAGAGAAATCAGCTGTGATGCTGAG 300
DB 411 CTGGAGCCAGTGAATACTCCCACTGGAGAGAGAGAGAGAAATCAGCTGTGATGCTGAG 470
QY 301 CCAAGAAATTCATTAAGATAGCAATCTGCTGCTTGGACAGATGGGATCTCTCCAGAA 360
DB 471 CCAAGAAATTCATTAAGATAGCAATCTGCTGCTTGGACAGATGGGATCTCTCCAGAA 530
QY 421 GAAGCAAGAGGGAATGTTGTTTATATACCAACCTTTCATCACTCACTCAAGACGGTG 480
DB 531 GAAGCAAGAGGGAATGTTGTTTATATACCAACCTTTCATCACTCACTCAAGACGGTG 590
QY 591 GAAGCAAGAGGGAATGTTGTTTATATACCAACCTTTCATCACTCACTCAAGACGGTG 650
DB 651 GAAGCAAGAGGGAATGTTGTTTATATACCAACCTTTCATCACTCACTCAAGACGGTG 710
QY 651 CAATACCGAAGCGCAGGGGGGGTGAAGCTGCGCAAGTGGGGGCTTTGGCATCTCTCAT 710
DB 711 CAATACCGAAGCGCAGGGGGGGTGAAGCTGCGCAAGTGGGGGCTTTGGCATCTCTCAT 770
QY 711 CGATCCGCGGCTCTCTTCCATCTACAGTCTCTCAACAGATTCAGAAATACAGAGAT 770
DB 771 CGATCCGCGGCTCTCTTCCATCTACAGTCTCTCAACAGATTCAGAAATACAGAGAT 830
QY 601 GGGCGGCCCAAAATTCACACAGCTGTATTAAGGAGAAATGATGTCAGAA 660
DB 771 GGGCGGCCCAAAATTCACACAGCTGTATTAAGGAGAAATGATGTCAGAA 830
QY 661 ATGGCTTCTCATGGGATCAAAATTTGCTCATTTCACTAAAGATGGGGGCAAAAGCTTACCA 720
DB 831 ATGGCTTCTCATGGGATCAAAATTTGCTCATTTCACTAAAGATGGGGGCAAAAGCTTACCA 890
QY 721 GATACCTGATTTCTTCAACACTGTAGACAGATCACTGGGAGCAATATTCAGAAAGATT 780
DB 891 GATACCTGATTTCTTCAACACTGTAGACAGATCACTGGGAGCAATATTCAGAAAGATT 950
QY 781 GTACGTGTCAGTGCATCTGACAGCTGGAGATGTTGGGACAGGGGCAATGATATGAC 840
DB 951 GTACGTGTCAGTGCATCTGACAGCTGGAGATGTTGGGACAGGGGCAATGATATGAC 1010
QY 841 GGTGAGGCTTTATATCATGAGAACACTCTCATTTATTAAGATCTTGGCTGGCTGCA 900
DB 1011 GGTGAGGCTTTATATCATGAGAACACTCTCATTTATTAAGATCTTGGCTGGCTGCA 1070
```

```

QY 901 AAGAGACCTGCGGGCTGTGCTCTGACTGACAGAAACAAGGTGAGTTGTCCTTC 960
DB 1071 AAGAGACCTGCGGGCTGTGCTCTGACTGACAGAAACAAGGTGAGTTGTCCTTC 1130
QY 961 CAGTATTATCAGTTACCAAGATTAATTTTCCAACTACAGTCTGTGATGAGTCTGAC 1020
DB 1131 CAGTATTATCAGTTACCAAGATTAATTTTCCAACTACAGTCTGTGATGAGTCTGAC 1190
QY 1021 GCAGGAACCTTTTACCCACTGGGCTGCAATTCCTGSCAGTGAAGAAAGCCAGGCCATC 1080
DB 1191 GCAGGAACCTTTTACCCACTGGGCTGCAATTCCTGSCAGTGAAGAAAGCCAGGCCATC 1250
QY 1081 ATGAGAGAGGTTATGAGCTGCTGACAGCCCTCAATATCACTGAGTCTGAGCCATGGA 1140
DB 1251 ATGAGAGAGGTTATGAGCTGCTGACAGCCCTCAATATCACTGAGTCTGAGCCATGGA 1310
QY 1141 GAAGGACAGACATCAACTTTTGGATCCAGCTGAGTGGCTGAGGCCAGTCTTATGAT 1200
DB 1311 GAAGGACAGACATCAACTTTTGGATCCAGCTGAGTGGCTGAGGCCAGTCTTATGAT 1370
QY 1201 GACTTATACAAATATTTCTTCTTCATCACTCCACGAGACACATGATCTGATGAT 1260
DB 1371 GACTTATACAAATATTTCTTCTTCATCACTCCACGAGACACATGATCTGATGAT 1430
QY 1261 CCAAGGAGATGATGTTGCTGCTGCTGTTGGCTGTTGGCTGTTTCTTATGTTTGCAGAC 1320
DB 1431 CCAAGGAGATGATGTTGCTGCTGCTGTTGGCTGTTGGCTGTTTCTTATGTTTGCAGAC 1490
QY 1321 ATGAGAGAAATGCTGCTAGGTCC 1344
DB 1491 ATGAGAGAAATGCTGCTAGGTCC 1514
```

RESULT 2

US-09-984-271-28
Sequence 28, Application US/0984271
Publication No. US20030040088A1

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins

FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/984,271
CURRENT FILING DATE: 2001-10-29

PRIOR APPLICATION NUMBER: 09/482,273
PRIOR FILING DATE: 2000-01-13

PRIOR APPLICATION NUMBER: PCT/US99/15849
PRIOR FILING DATE: 1999-07-14

PRIOR APPLICATION NUMBER: 60/092,921
PRIOR FILING DATE: 1998-07-15

PRIOR APPLICATION NUMBER: 60/092,922
PRIOR FILING DATE: 1998-07-15

PRIOR APPLICATION NUMBER: 60/092,956
PRIOR FILING DATE: 1998-07-15

NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 28

LENGTH: 1863

TYPE: DNA
ORGANISM: Homo sapiens

US-09-984-271-28
Query Match 99.8%; Score 1340.8; DB 11; Length 1863;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AAGATGGCATCTCTAAGAGACCTTTGAAGAAATGAAGAAATGCGGCTGTGGA 60
DB 171 AAGATGGCATCTCTAAGAGACCTTTTGAAGAAATGAAGAAATGCGGCTGTGGA 230
QY 61 GATGTTGCTAAAGCATCATCACTAGCTGTTTATGGTAAAGCCGAGAACAGATCTAT 120
DB 231 GATGTTGCTAAAGCATCATCACTAGCTGTTTATGGTAAAGCCGAGAACAGATCTAT 290
```

```

QY 121 GAGCGATTGGCACTTCTGTTGATGATGTTGACCCGAGACTGAGTGGCTCCAGAACTTA 180
DB 291 GAGCGATTGGCACTTCTGTTGATGATGTTGACCCGAGACTGAGTGGCTCCAGAACTTA 350
QY 181 GAAAAAGCCATCCCAATTATGTATGACCAAAACCTGAGAGAGTGGCTGAGAAAGTTTCA 240
DB 351 GAAAAAGCCATCCCAATTATGTATGACCAAAACCTGAGAGAGTGGCTGAGAAAGTTTCA 410
QY 241 CTGAGCGAGTGAAGTAAATCCCACTGGGAGAGGGGAGAAAGAAATGAGTGTGAGCTGAG 300
DB 411 CTGAGCGAGTGAAGTAAATCCCACTGGGAGAGGGGAGAAAGAAATGAGTGTGAGCTGAG 470
QY 301 CCAAGATTCATTAAGATAGCCATCTGGGATCTTGGCAGAGCATTTGGGACTCTCCAGAA 360
DB 471 CCAAGATTCATTAAGATAGCCATCTGGGATCTTGGCAGAGCATTTGGGACTCTCCAGAA 530
QY 361 GGCATTACAGCAGAAAGTTTGTGTGTGATGACCTCTTTCATGATGACAGAGAGGGCTCA 420
DB 531 GGCATTACAGCAGAAAGTTTGTGTGTGATGACCTCTTTCATGATGACAGAGAGGGCTCA 590
QY 421 GAAGCAAGAGGAGATGTTGTTTATTAACCACTTAATCACTACTCAAGAGCGGTG 480
DB 591 GAAGCAAGAGGAGATGTTGTTTATTAACCACTTAATCACTACTCAAGAGCGGTG 650
QY 481 CAATACCGAAGCGAGGGGCGGTGGAAGCTGCCAAGGTGGGGCTTTGGCATCTCTCAT 540
DB 651 CAATACCGAAGCGAGGGGCGGTGGAAGCTGCCAAGGTGGGGCTTTGGCATCTCTCAT 710
QY 541 CGATCCGTGGCTCTCTCTCATCTACAGTCTTCAACAGGATTTCAAGAAATACAGAGAT 600
DB 711 CGATCCGTGGCTCTCTCTCATCTACAGTCTTCAACAGGATTTCAAGAAATACAGAGAT 770
QY 601 GGGGTGCGCCAAATTCACAGCGCTGTATTAACGTGTGAAGATGCAAAATGATCAAGA 660
DB 771 GGGGTGCGCCAAATTCACAGCGCTGTATTAACGTGTGAAGATGCAAAATGATCAAGA 830
QY 661 ATGGCTCTCATGGGATCAAAATTTGTCATTCAGTTAAAGATGGGGGCAAGACTTACCA 720
DB 831 ATGGCTCTCATGGGATCAAAATTTGTCATTCAGTTAAAGATGGGGGCAAGACTTACCA 890
QY 721 GATACGTATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAACAGATT 780
DB 891 GATACGTATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATTCAGAAACAGATT 950
QY 781 GTACTGTGCTGAGTCACTGTGACAGCTGGGATGTTGGGCGAGGCTGCATGATGATGCG 840
DB 951 GTACTGTGCTGAGTCACTGTGACAGCTGGGATGTTGGGCGAGGCTGCATGATGATGCG 1010
QY 841 GGTGAGCGCTTATATATATGAGGAGCACTCTCACTTAATTAAGTCTTGGGCTGGCTCA 900
DB 1011 GGTGAGCGCTTATATATATGAGGAGCACTCTCACTTAATTAAGTCTTGGGCTGGCTCA 1070
QY 901 AAGAGCACTCTGCGGCTGTGCTCTGAGCTGACAGAGAAACAGAGTGGATGTCCTTC 960
DB 1071 AAGAGCACTCTGCGGCTGTGCTCTGAGCTGACAGAGAAACAGAGTGGATGTCCTTC 1130
QY 961 CAGTATTATGATTAACAAGATTAATTTCCAACTACAGTCTGTGATGATGATGATGAC 1020
DB 1131 CAGTATTATGATTAACAAGATTAATTTCCAACTACAGTCTGTGATGATGATGATGAC 1190
QY 1021 GAGAGAACTTCTTAACCACTGGGCTCAANTCACTGCGAGTGAAGAAAGCCAGGGCCATC 1080
DB 1191 GAGAGAACTTCTTAACCACTGGGCTCAANTCACTGCGAGTGAAGAAAGCCAGGGCCATC 1250
QY 1081 ATGAGAGAGGTTATGAGCTGCTGACGCCCTCAATATCACTCAGTCTTGAGCCATGGA 1140
DB 1251 ATGAGAGAGGTTATGAGCTGCTGACGCCCTCAATATCACTCAGTCTTGAGCCATGGA 1310
QY 1141 GAAGGAGCAGATCAACTTTTGATTCAGAGCTGAGTGGCTGGAGCCAGTCTTACTTGAT 1200
DB 1311 GAAGGAGCAGATCAACTTTTGATTCAGAGCTGAGTGGCTGGAGCCAGTCTTACTTGAT 1370

```

```

QY 1201 GACTTATACAGATATTTCTTCTTCATCACTCCCAAGAGACACATGACTGTATGAT 1260
DB 1371 GACTTATACAGATATTTCTTCTTCATCACTCCCAAGAGACACATGACTGTATGAT 1430
QY 1261 CCAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1320
DB 1431 CCAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1490
QY 1321 ATGAGAGAAATGCTGCCTAGGTCC 1344
DB 1491 ATGAGAGAAATGCTGCCTAGGTCC 1514

RESULT 3
US-09-917-800A-505
; Sequence 505, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917, 800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 505
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURES:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AF097723
US-09-917-800A-505

Query Match 78.9%; Score 1060.8; DB 10; Length 1778;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1167; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 1 AAGATGGCATCTCTAAGAGCACTTTGAAGAAATTAAGAAATAGCAGCTGGA 60
DB 193 AAGATGGCTTTCTCAGCCAACTTTCAGAAATTAAGAAATAGCAGCTGGA 252
QY 61 GATGTTGCTAAGCAATCATCAACTAGCTGTTATGTTAAGCCAGAAACAGATCTAT 120
DB 253 GATGTTGCTAAGCAATCATCAACTAGCTGTTATGTTAAGAAATTAAGCCAGATCTAT 312
QY 121 GAGCGATTGGCACTTCTGTTGATGATGTTGACCCGAGACTGAGTGGCTCCAGAACTTA 180
DB 313 GAGCGATTGGCACTTCTGTTGATGATGTTGACCCGAGACTGAGTGGCTCCAGAACTTA 372
QY 181 GAAAAAGCCATCCCAATTATGTATGACCAAAACCTGAGAGAGTGGCTGAGAAAGTTTCA 240

```

Db 373 GAGAAAGTATCCAAATCATGTACCAAAACCTGCAACAAGATGGCTGAAAAGCTTCAC 432
Qy 241 CTGAGCCAGTGAAGATCCCACTGGAGAGGGGAGAAATCACTGTGTGTGGAG 300
Db 433 CTGGAGCGGTGAGAAATCTCACTGGGGGAGGGGAGAAATCTGCAGTATGTGTGTG 492
Qy 301 CCAAGAATTCATPAAGATTCCTGAGTCTGGCAGACAGATGGAGACTCTCCAGAA 360
Db 493 CTTGAATTCACAAATTCCTGATTTTAAAGCTTGGCGAGCATTTGGAGCTCTCTGAA 552
Qy 361 GGCATTACAGCAGAGATCTGGTGTGACTCTTTGATGAATCTGCAGAGAGGGCTCA 420
Db 553 GGTATCACAGCAGAGATCTGGTGTGACTCTTTGATGAATCTCAAGAGGGCATCA 612
Qy 421 GAAGCAGAGAGAGATTTGTTTATACCACTTACATCACTACTCAAGAGGGTG 480
Db 613 GAGGCAGAGAGAGATTTGTTTATACCACTTACATCACTACTCAAGAGGGTG 672
Qy 481 CAATACCGAAGCGAGGGGGCGGTGAGCTGCAGAGTGGGGGCTTTGGCATCTCTCAT 540
Db 673 CAGTACCGGAGCGGAGCTGTGAGCTGCAGAGTGGGGGCGGTGAGCTCTCTCATC 732
Qy 541 CGATCCGTGGCTCTCTCTCATCTACAGTCTCTCAAGAGTATTCAGAGATACAGAGAT 600
Db 733 CGATAGTAGTCTCTTTTTCATCTACAGTCTCTCAAGAGTATTCAGAGATATCAAGAT 792
Qy 601 GCGGTGCCCAAAATTCACACAGCTGTATTAACGTTGAGAGATGCAGAAATGATCAAGA 660
Db 793 GGTGTGCCCAAGATTCACACAGCTGTATTAACGTTGAGAGATGCAGAAATGATCTCGA 852
Qy 661 ATGGCTTCATGAGGATCAAAATTTGTCATTCAGTCAAGATGGGGGCGAAAGCCATCCA 720
Db 853 ATGGCTTCATGAGGATCAAAATTTGTCATTCAGTCAAGATGGGGGCGAAAGCCATCCA 912
Qy 721 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGT 780
Db 913 GATACAGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGT 972
Qy 781 GTACTGTGTCAGTGCACATCTGGACAGCTGGGATGTTGGGCGAGGTGCCATGATGATGC 840
Db 973 GTCTGTGTGTCAGTGCACATCTGGACAGCTGGGATGTTGGGCGAGGTGCCATGATGATGC 1032
Qy 841 GGTGAGCTTTATATCATGGGAGACATCTCACTATTAAAGATCTTGGGCTGTGTCA 900
Db 1033 GGTGAGCTTTATATCATGGGAGACATCTCACTATTAAAGATCTTGGGCTGTGTCA 1092
Qy 901 AAGAGACTCTGCGGCTGTGTCTGGACTGCAGAGAGAAACAAGTGAATGTGTGCTTC 960
Db 1093 AAGAGACTCTGCGGCTGTGTCTGGACTGCAGAGAGAAACAAGTGAATGTGTGCTTC 1152
Qy 961 CAGTATTCAGTTCACAGAGTAAATTTTCAACTCAAGTCTGTGTATGAGATCTGAC 1020
Db 1153 CAGTATTCAGTTCACAGAGTAAATTTTCAACTCAAGTCTGTGTATGAGATCTGAC 1212
Qy 1021 GCAGGAACTCTTACCACTGGGCTGGAATTCATGSCAGAGTGAAGAGGAGGAGGCTGC 1080
Db 1213 TCAGGAACTCTTACCACTGGGCTGGAATTCATGSCAGAGTGAAGAGGAGGAGGCTGC 1272
Qy 1081 ATGAGAGGATTAAGAGCTGTGAGGCTGCAGGCTCAATATCACTCAAGTCTGAGGCAATGA 1140
Db 1273 ATGAGAGGATTAAGAGCTGTGAGGCTGCAGGCTCAATATCACTCAAGTCTGAGGCAATGA 1332
Qy 1141 GAAGGAGACAGATCACTTTTGGATCCAGCTGAGTGTCTGAGGAGGCTGTACTTGAT 1200
Db 1333 GAAGGAGACAGATCACTTTTGGATCCAGCTGAGTGTCTGAGGAGGCTGTACTTGAT 1392
Qy 1201 GACTTATCAAGTATTTCTTCCATCACTCCCAAGGAGCAACATGATCTGATGAT 1260
Db 1393 GACTTATCAAGTATTTCTTTTCCATCACTCCCAAGGAGCAACATGATCTGATGAT 1452
Qy 1261 CCAAGAGATGATGTTGCTGTGCTGTGAGGCTGTGTTCTTATGTTGACAGC 1320

Db 1453 CAAAGCAGATGATGTGCTGTGCTGTTTGGGCTGTTCGCTTACGTTGGCAGAC 1512
Qy 1321 ATGAGAGAAATGCTGCTAGGTCC 1344
Db 1513 ATGAGAGAAATGCTGCCAGGTCC 1536

RESULT 4
US-09-984-271-95
; Sequence 95, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-271-95

Query Match 57.5%; Score 772.4; DB 11; Length 1134;
Best Local Similarity 99.7%; Pred. No. 4.9e-241;
Matches 784; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 559 TCATCTACAGTCTCTCAACAGGATTCAGAGATCCAGAGATGGCGCCCAAAATTC 618
Db 1 TCATCTACAGTCTCTCAACAGGATTCAGAGATCCAGAGATGGCGCCCAAAATTC 60
Qy 619 ACAGCTGTATTCAGGAGAGATGCAGAAATGATCAAGATGGCTTCATGGGATC 678
Db 61 ACAGCTGTATTCAGGAGAGATGCAGAAATGATCAAGATGGCTTCATGGGATC 120
Qy 679 AAAATGTCAATTCAGCTAAAGATGGGGCAAGACTTACCAGATTCCTTCAAC 738
Db 121 AAAATGTCAATTCAGCTAAAGATGGGGCAAGACTTACCAGATTCCTTCAAC 180
Qy 739 ACTGTACAGAGATCACTGGGAGCAATATTCAGAAACAGGTTGTACTGTGTGAGCAT 798
Db 181 ACTGTACAGAGATCACTGGGAGCAATATTCAGAAACAGGTTGTACTGTGTGAGCAT 240
Qy 799 CTGAGACAGTGGAGTGTGGGCAAGGTCATGAGTGAATGAGGCTTGAATTC 858
Db 241 CTGAGACAGTGGAGTGTGGGCAAGGTCATGAGTGAATGAGGCTTGAATTC 300
Qy 859 TGGAGACACTTCATCTTAAAGATCTTGGGCTGCTCAAGAGACTTGGGCTG 918
Db 301 TGGAGACACTTCATCTTAAAGATCTTGGGCTGCTCAAGAGACTTGGGCTG 360
Qy 919 GTGCTGTGAGTGCAGAGAAACAAGGTGTGTGTCTTCCAGTATTCAGTTCAC 978
Db 361 GTGCTGTGAGTGCAGAGAAACAAGGTGTGTGTCTTCCAGTATTCAGTTCAC 420
Qy 979 AAGGTAAATTTTCCATCAAGCTGTGTGATGAGTCTGAGAGAACTTTTACC 1038
Db 421 AAGGTAAATTTTCCATCAAGCTGTGTGATGAGTCTGAGAGAACTTTTACC 480
Qy 1039 ACTGGGTGAATTCATGTGCAAGTGAAGAGGCTCAATGAGAGGTTATGAGC 1098

Db 481 ACTGGGCTGCATTCCTGCGAGTGAAAAGCCAGGG-CATCATGAGAGGTTATGAC 539
Qy 1099 CTGCTGAGAGCCCTCAATATCACTCAGTCTCTGAGCAGTGAAGGAGACATCAAC 1158
Db 540 CTGCTGAGAGCCCTCAATATCACTCAGTCTCTGAGCAGTGAAGGAGACATCAAC 599
Qy 1159 TTTTGGATCCAGAGTGAAGTCCCTGAGAGCCAGTCTACTTGAATATACAGTATTC 1218
Db 600 TTTTGGATCCAGAGTGAAGTCCCTGAGAGCCAGTCTACTTGAATATACAGTATTC 659
Qy 1219 TTTTCCATCATCTCCAGAGGAGACATGATCTGATGATTCGAAAGCAGTAAATTT 1278
Db 660 TTTTCCATCATCTCCAGAGGAGACATGATCTGATGATTCGAAAGCAGTAAATTT 719
Qy 1279 GCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTTGAGACATGAGAAATGCTGCT 1338
Db 720 GCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTTGAGACATGAGAAATGCTGCT 779
Qy 1339 AGGTCC 1344
Db 780 AGGTCC 785

RESULT 5
US-09-833-381-1929
; Sequence 1929, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833.381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1929
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(357)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1929

Query Match 19.5%; Score 261.8; DB 10; Length 357;
Best Local Similarity 86.4%; Pred. No. 1.6e-74;
Matches 298; Conservative 0; Mismatches 46; Indels 1; Gaps 1;

Qy 898 CCAAAGAGAGCTTCGCGGCTGCTGCTGAGCTGAGAAAGCAAGGTGGTGGC 957
Db 13 CGAAGAGAGCTTCGCGGCTGCTGCTGAGCTGAGAAAGCAAGGTGGTGGC 72
Qy 958 TTCAGATTATCACTTACCAAGGTAATTTTCCACTGACGTGGTGAATGAGTCT 1017
Db 73 TCCAGATTATCACTTACCAAGGTAATTTTCCACTGACGTGGTGAATGAGTCT 132
Qy 1018 GAGCAGAGAGCTTCTTACCACTGGGCTGCAATTCATGCGAGTGAAGAGCCAGGCGC 1077
Db 133 GAGCAGAGAGCTTCTTACCACTGGGCTGCAATTCATGCGAGTGAAGAGCCAGGCGC 192
Qy 1078 ATCATGAGAGGTTATGAGCTGCTGAGAGCCCTCAATATCACTCAGTCTTGAACCAT 1137
Db 193 ATCATGAGAGGTTATGAGCTGCTGAGAGCCCTCAATATCACTCAGTCTTGAAT 252
Qy 1138 GAGAGAGGAGAGCATGATTTTGAATGCAAGC-TGAGATGCTGGAGGAGTCACT 1196
Db 253 GAGAGAGGAGAGCATGATTTTGAATGCAAGC-TGAGATGCTGGAGGAGTCACT 312
Qy 1197 TGATGACTTATCAAGTATTTCTTTCATCACTCCACGAGAG 1241

Db 313 TGATGACTTATCAAGTATTTCTTTCATCACTCCACGAGAG 357

RESULT 6
US-09-833-381-1930
; Sequence 1930, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833.381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1930
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(427)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1930

Query Match 16.1%; Score 215.8; DB 10; Length 427;
Best Local Similarity 87.7%; Pred. No. 1.9e-59;
Matches 235; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 AAGATGGCATCTCTAAGAGAGCTTTGAAGATATAAAGAGAAATAGCCAGTGTGA 60
Db 157 AAGATGGCTTTCTCAGCGAACAATTTGAGAAATATAAGAGAAATAGCCACTATGA 216
Qy 61 GATGTTCTAAGCAATCAATCAACTAGCTGTTATGTTAAAGCCAGAACATCTAT 120
Db 217 GATGTTCTAAGCAATCAATCAACTAGCTGTTATGTTAAAGCCAGAACATCTAT 276
Qy 121 GAGCGATTGCACTTCTGTTGATGATCTGTTGAGACCCAGACTGAGTGGCTCAAGAACTTA 180
Db 277 GAGCGATTGCACTTCTGTTGATGATCTGTTGAGACCCAGACTGAGTGGCTCAAGAACTTA 336
Qy 181 GAAAGAGCCATCAATTAATGATGCAAAACCTGACAGAGAGTGGCTGAGAAATGTCAC 240
Db 337 GAGAAAGCTATTCAATCATGATGCAAAACCTGACAGAGAGTGGCTGAGAAATGTCAC 396
Qy 241 CTGAGCCAGTGAATATCCCACTGGG 268
Db 397 CTGAGCCAGTGAATATCCCACTGGG 424

RESULT 7
US-09-783-590-12141
; Sequence 12141, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783.590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12141

PRIOR APPLICATION NUMBER: US 60/183,322
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: US 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 11-248036
PRIOR FILING DATE: 1999-07-29
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (173)...(1450)
US-10-059-585-5

Query Match 3.1%; Score 42; DB 15; Length 2200;
Best Local Similarity 45.5%; Pred. No. 0.033;
Matches 150; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 763 AATATCCAGAACAGTTGTACTGTCAGATCGATCGAGACCTGGAGTGGCAG 822
DB 158 AAGATGAGAGAGGATGCTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 217
QY 823 GTGCCATGATGATGAGCGGTGAGAGCTTTATATATGAGAGAGAGAGAG 882
DB 218 GACAGCCCTGACAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 277
QY 883 GATCTTGGGCTGCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 942
DB 278 GAG 337
QY 943 GGTGAG 1002
DB 338 GAG 397
QY 1003 CTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1062
DB 398 ATCCCTGCTCTTATATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 457
QY 1063 GAAAG 1092
DB 458 GAG 487

RESULT 11
US-09-908-975-1202
Sequence 1202, Application US/09908975
Publication No. US20030165843A1
GENERAL INFORMATION:
APPLICANT: SHOSHAN, Avi
APPLICANT: MASSERMAN, Alon
APPLICANT: MINTZ, Eli
APPLICANT: FAIGLER, Simchon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
FILE REFERENCE: 36688-0005
CURRENT APPLICATION NUMBER: US/09/908,975
CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1202
LENGTH: 65

TYPE: DNA
ORGANISM: Rattus norvegicus
US-09-908-975-1202

Query Match 3.1%; Score 41; DB 13; Length 65;
Best Local Similarity 76.9%; Pred. No. 0.0078;
Matches 50; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 271 AGGGAG 330
DB 1 AGGGAG 60
QY 331 CTGG 335
DB 61 CTGG 65

RESULT 12
US-09-918-995-30889

Sequence 30889, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:

APPLICANT: Hyeq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30889

LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(484)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30889

Query Match 2.9%; Score 39.4; DB 11; Length 484;
Best Local Similarity 49.3%; Pred. No. 0.091;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 144 TACTGTGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 203
DB 231 TAGTGTCTGTTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 290
QY 204 CCAAACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 263
DB 291 CCTGAGAGATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 350
QY 264 CTGGAG 323
DB 351 AGGGAG 410
QY 324 CTGGGCTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 352
DB 411 CAGAGATCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 439

RESULT 13
US-10-232-484-5

Sequence 5, Application US/10232484
Publication No. US20030113847A1
GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Yonghong
APPLICANT: Drmanac, Radoje T
TITLE OF INVENTION: No. US20030113847A1 Nucleic Acids and Polypeptides
FILE REFERENCE: 791CIP2BDIV

CURRENT APPLICATION NUMBER: US/10/232,484
CURRENT FILING DATE: 2002-08-30
PRIOR APPLICATION NUMBER: 09/695,783
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: 09/552,929
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1363
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (848)..(1090)
OTHER INFORMATION:
US-10-232-484-5

Query Match 2.9%; Score 39.4; DB 15; Length 1363;
Best Local Similarity 49.3%; Pred. No. 0.17;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 144 TACTGTGGACCCGACAGTGTGCTCCAGAACTAGAAAAAGCCATCAATTATGTA 203
DB 877 TAGTGTCTGTTCCAAACGCACTGTGTCAAGAACTAGAAATTAACATTAGAGGCAAAAC 936
QY 204 CCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAGAGCCAGTGAATACCCCA 263
DB 937 CCTGAGAAATACACAAGGGGGGACCTTCCAGTAAGTGTGTTGGGAAAGAGAGAGGAG 996
QY 264 CTGGAGAGGGGAGAGAAATCAGTGTGATCTGAGAGCCAAAGATTCAATAGATAGCCAT 323
DB 997 AGGGGACAGGGGAGACAGATTCAGCTTTGTGTGGGTCCCTGAGGGTTCTTACAGGGGTAG 1056
QY 324 CCTGGCTCTTGGCAGCAGCATTTGGGACTC 352
DB 1057 CCAGGATCTGGGAAACAGATCAGGCACTC 1085

RESULT 14
US-10-044-090-579
Sequence 579, Application US/10044090
Publication No. US20020137081A1
GENERAL INFORMATION:
APPLICANT: Olga Bandman
TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
FILE REFERENCE: PA-0028 US
CURRENT APPLICATION NUMBER: US/10/044,090
CURRENT FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 850
SOFTWARE: PERL Program
SEQ ID NO 579
LENGTH: 2319
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20020137081A1 023244.1
LOCATION: 169, 173
OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-579

Query Match 2.9%; Score 39.4; DB 14; Length 2319;
Best Local Similarity 49.3%; Pred. No. 0.24;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 144 TACTGTGGACCCGACAGTGTGCTCCAGAACTAGAAAAAGCCATCAATTATGTA 203
DB 1560 TAGTGTCTGTTCCAAACGCACTGTGTCAAGAACTAGAAATTAACATTAGAGGCAAAAC 1619
QY 204 CCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAGAGCCAGTGAATACCCCA 263

DB 1620 CCTGAGAAATACACAAGGGGCAAGCTTCCAGTATGTGTGGGAGAGAGAGGAGCAG 1679
QY 264 CTGGAGAGGGGAGAGAAATCAGTGTGATCTGAGAGCCAAAGATTCAATAGATAGCCAT 323
DB 1680 AGGGGACAGGGGAGACAGATTCAGCTTTGTGTGGGTCCCTGAGGGTTCTTACAGGGGTAG 1739
QY 324 CCTGGCTCTTGGCAGCAGCATTTGGGACTC 352
DB 1740 CCAGGATCTGGGAAACAGATCAGGCACTC 1768

RESULT 15
US-09-814-353-20277
Sequence 20277, Application US/09814353
Publication No. US20030165831A1
GENERAL INFORMATION:
APPLICANT: Lee, John
APPLICANT: Thompson, Pamela
APPLICANT: Lillie, James
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
FILE REFERENCE: MRI-006B
CURRENT APPLICATION NUMBER: US/09/814,353
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,820
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 20277
LENGTH: 3322
TYPE: DNA
ORGANISM: Homo sapiens
US-09-814-353-20277

Query Match 2.9%; Score 39.4; DB 13; Length 3322;
Best Local Similarity 49.3%; Pred. No. 0.3;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 144 TACTGTGGACCCGACAGTGTGCTCCAGAACTAGAAAAAGCCATCAATTATGTA 203
DB 1946 TAGTGTCTGTTCCAAACGCACTGTGTCAAGAACTAGAAATTAACATTAGAGGCAAAAC 2005
QY 204 CCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAGAGCCAGTGAATACCCCA 263
DB 2006 CCTGAGAAATACACAAGGGGCAAGCTTCCAGTATGTGTGGGAAAGAGAGGAGCAG 2065
QY 264 CTGGAGAGGGGAGAGAAATCAGTGTGATCTGAGAGCCAAAGATTCAATAGATAGCCAT 323
DB 2066 AGGGGACAGGGGAGACAGATTCAGCTTTGTGTGGGTCCCTGAGGGTTCTTACAGGGGTAG 2125
QY 324 CCTGGCTCTTGGCAGCAGCATTTGGGACTC 352
DB 2126 CCAGGATCTGGGAAACAGATCAGGCACTC 2154

Search completed: December 22, 2003, 23:44:41
Job time: 421.737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:18:08 ; Search time 2046.81 Seconds

(without alignments)
11332.658 Million cell updates/sec

Title: US-09-745-763-35_COPY_57_623

Perfect score: 567

Sequence: 1 CCTATCAGATTATCTTAAACA.....AGATTGTTGTTATTAACCAA 567

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapept 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl1.*

1: gb Da.*

2: gb Htg.*

3: gb In.*

4: gb Cm.*

5: gb Ov.*

6: gb Pat.*

7: gb Ph.*

8: gb Pl.*

9: gb Pr.*

10: gb Ro.*

11: gb Sts.*

12: gb Sy.*

13: gb Un.*

14: gb Vi.*

15: em Ba.*

16: em Fun.*

17: em Hum.*

18: em In.*

19: em Mu.*

20: em Om.*

21: em Or.*

22: em Ov.*

23: em Pat.*

24: em Ph.*

25: em Pl.*

26: em Ro.*

27: em Sts.*

28: em Un.*

29: em Vi.*

30: em Htg Hum.*

31: em Htg Inv.*

32: em Htg Other.*

33: em Htg Mus.*

34: em Htg Pln.*

35: em Htg Rod.*

36: em Htg Mam.*

37: em Htg Vrt.*

38: em Hy.*

39: em Htgo Hum.*

40: em Htgo Mus.*

41: em Htgo Other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	567	100.0	1472	9 AF107834	AF107834 Homo sapi
2	567	100.0	1851	6 BD106411	BD106411 Secreted
3	567	100.0	1860	6 AX006440	AX006440 Sequence
4	567	100.0	1928	6 BC020689	BC020689 Homo sapi
5	563.8	99.4	1860	6 BD127520	BD127520 Primer fo
6	563.8	99.4	1860	6 AK075132	AK075132 Homo sapi
7	555	97.9	609	6 BD059610	BD059610 Secreted
8	549	96.8	1794	9 AF119386	AF119386 Homo sapi
9	526.2	92.8	1778	6 BD105816	BD105816 Secretory
10	512.2	90.3	642	6 BD125210	BD125210 Primer fo
11	512.2	90.3	642	6 BD126476	BD126476 Primer fo
12	472.4	83.3	156702	2 AC084223	AC084223 Homo sapi
13	472.4	83.3	157927	2 AP006278	AP006278 Homo sapi
14	472.4	83.3	171301	9 AC010859	AC010859 Homo sapi
15	400	70.5	1726	10 AF131077	AF131077 Rattus no
16	399.6	70.5	1778	6 AX400829	AX400829 Sequence
17	399.6	70.5	1778	10 AF097723	AF097723 Rattus no
18	395.6	69.8	1716	10 AF009513	AF009513 Mus muscu
19	395.6	69.8	1796	10 BC037067	BC037067 Mus muscu
20	394	69.5	1596	10 AF107835	AF107835 Mus muscu
21	327.4	57.7	167883	2 AC121026	AC121026 Rattus no
22	327.4	57.7	218874	2 AC110103	AC110103 Rattus no
23	323.6	57.1	214287	2 AC129951	AC129951 Mus muscu
24	296	52.2	423	9 AF107833	AF107833 Homo sapi
25	290.8	51.3	314	6 BD058492	BD058492 Secreted
26	284.6	50.2	462	6 BD105811	BD105811 Secretory
27	284.6	50.2	462	6 E62849	E62849 Secretary P
28	254.6	44.9	317	6 BD076934	BD076934 5' EST of
29	121	21.3	217	6 BD076930	BD076930 5' EST of
30	49.2	8.7	7218	6 I66494	I66494 Sequence 14
31	46.6	8.2	1729	3 AF077194	AF077194 Acanthoch
32	45.2	8.0	161084	2 AC024030	AC024030 Homo sapi
33	45.2	8.0	161488	9 AP003083	AP003083 Homo sapi
34	45.2	8.0	173003	2 AC024245	AC024245 Homo sapi
35	41.8	7.4	259599	2 AC098081	AC098081 Rattus no
36	41.8	7.4	262687	2 AC128747	AC128747 Rattus no
37	41.6	7.3	230426	9 AC017099	AC017099 Homo sapi
38	41.2	7.3	266459	2 AC099884	AC099884 Mus muscu
39	41	7.2	3909	1 BA011045	BA011045 Buchnera ap
40	41	7.2	11249	1 AB014105	AB014105 Buchnera
41	40.6	7.2	10792	1 AB015790	AB015790 Shewanella
42	40.6	7.2	183155	2 AC128618	AC128618 Rattus no
43	40.6	7.2	253140	2 AC094055	AC094055 Rattus no
44	40.2	7.1	110000	2 FPMAL6P1_05	Continuation (6 of
45	40.2	7.1	162065	2 AC117061	AC117061 Rattus no

ALIGNMENTS

RESULT 1

AF107834

LOCUS AF107834 1472 bp mRNA linear PRI 12-JUL-1999

DEFINITION Homo sapiens clone LCHI aminopeptidase mRNA, complete cds.

ACCESSION AF107834

VERSION AF107834.1 GI:5442029

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

Homosapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1472)

AUTHORS Liu,C.H., Lan,B.Y. and Chang,L.Y.

TITLE Cloning of the human aminopeptidase gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1472)
 AUTHORS Liu, C. H., Lin, B. Y. and Chang, L. Y.
 TITLE Direct Submission
 JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia Sinica, RM 433, 128, Yen-Chun-Yuan Road SEC 2, Taipei 11529, Taiwan

FEATURES
 source
 1. 1472
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="Uchi"
 /tissue_type="liver"
 /dev_stage="fetus"
 54. 1472
 /note="secretory protein"
 /codon_start=1
 /product="aminopeptidase"
 /protein_id="AAD43214.1"
 /db_xref="GI:5442030"
 /translation="MKFLIFAFPGVHLTSLCSGKAIKNGISKRTPEEIKETIASGCS
 DVAKRIINLAVYGRKONSVERLALINDTVPRRSGENLEKAQIMQNIQDGLK
 VHAEPVRIPIHMERGESAVMLEPPIHKIAIIGLSSIGTPPEGITAEVLVTSDELQ
 RRAEARKIVINQPIYINSTRVQYRTQGAVEAAKVALASLRVASFISYSPHTG
 IQEYQDGVPIPTACITVEDAMMSRMSHGKIKIVIQKMAKTYPTDSFNTVAEIT
 GSKYPEGVLVSGHLSMDVQCAMDGGAFISMEALSLIKDGLRPRKTLRLVLT
 AEEQGVGAFOYQOLHKVNISNLSLVMSDAGTLPGLTPTGSGSEKRAIMEEVMSLL
 QPLNITQLSHGEGDINPMIOAGVPRGSLDDLKRYFFHSHSGDMTVDPKQMV
 AAVVAIVSYVAAMEMLPRS"

CDS
 411 a 325 c 381 g 355 t

BASE COUNT
 ORIGIN

Query Match
 Best Local Similarity 100.0%; Score 567; DB 9; Length 1472;
 Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCTATCAGATTATCTTAAACAAGAAAACCACTGGAAAAAAATGAATTCCTTATCTTC
 60
 12 CCTATCAGATTATCTTAAACAAGAAAACCACTGGAAAAAAATGAATTCCTTATCTTC
 71
 61 GCATTTTCGGTGTGTTACCTTTTATCCTGTGCTCTGGAAAAGCTATTCGAAGAT
 120
 72 GCATTTTCGGTGTGTTACCTTTTATCCTGTGCTCTGGAAAAGCTATTCGAAGAT
 131
 121 GCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGT
 180
 132 GCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGT
 191
 181 GCTAAAGCAATCATCAACTGCTGTTTATGTATGTAAGCCAGAACAGATCTTATGAGCA
 240
 192 GCTAAAGCAATCATCAACTGCTGTTTATGTATGTAAGCCAGAACAGATCTTATGAGCA
 251
 241 TTGGCACTTCTGTGTGATCTGTTGAGCCGACGACTGAGTGGCTCCAGAACCTTGA
 300
 252 TTGGCACTTCTGTGTGATCTGTTGAGCCGACGACTGAGTGGCTCCAGAACCTTGA
 311
 301 GCCATCCAAATTATGTACCAAAAACCTGCGAGCAAGTGGGCTGAGAAAGTTCACTGGAG
 360
 312 GCCATCCAAATTATGTACCAAAAACCTGCGAGCAAGTGGGCTGAGAAAGTTCACTGGAG
 371
 361 CCACTGAGAAATACCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGACCA
 420
 372 CCACTGAGAAATACCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGACCA
 431
 421 ATTATATAGATAGCATCTCTGGGCTTGGCGACGACATTTGGGACTCTCTCCAGAAAGCA
 480
 432 ATTATATAGATAGCATCTCTGGGCTTGGCGACGACATTTGGGACTCTCTCCAGAAAGCA
 491
 481 ACAGCAGAAGTTCTGTGTGATGACCTCTTGTGATGAACTGAGAGAAAGGCTCTGAGAA
 540
 492 ACAGCAGAAGTTCTGTGTGATGACCTCTTGTGATGAACTGAGAGAAAGGCTCTGAGAA
 551
 541 AGAGGAGAGATTTGTTTATTAACCA 567

Db 552 AGAGGAGAGATTTGTTTATTAACCA 578

RESULT 2
 LOCUS BD106411 1851 bp DNA linear PAT 18-SEP-2002
 DEFINITION Secreted proteolins and polynucleotides encoding them.
 ACCESSION BD106411
 VERSION BD106411.1 GI:23201229
 KEYWORDS JP 2002503955-A/2.
 SOURCE
 ORGANISM
 Chlamydia sp.
 Bacteria; sp.
 Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

REFERENCE
 1 (bases 1 to 1851)
 AUTHORS Jacobs, K., McCoy, J. M., Lavallie, E. R., Racie, L. A., Merberg, D.,
 Treacy, M., Spaulding, V., and Agostino, M. J.
 TITLE Secreted proteins and polynucleotides encoding them
 JOURNAL Patent: JP 2002503955-A/2 05-FEB-2002;
 GENETICS INSTITUTE INC
 PN JP 2002503955-A/2
 PD 05-FEB-2002
 PF 20-MAR-1998 JP 1998545874
 PR 21-MAR-1997 US 08/822167, 19-MAR-1998 US 09/044466 PI
 KENNETH JACOB, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
 DAVID MERBERG,
 PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
 C12N15/12, C07K14/47, A61K38/17
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key

FEATURES
 source
 1. 1851
 /organism="Chlamydia sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:35827"

BASE COUNT
 ORIGIN

Query Match
 Best Local Similarity 100.0%; Score 567; DB 6; Length 1851;
 Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCTATCAGATTATCTTAAACAAGAAAACCACTGGAAAAAAATGAATTCCTTATCTTC
 60
 57 CCTATCAGATTATCTTAAACAAGAAAACCACTGGAAAAAAATGAATTCCTTATCTTC
 116
 61 GCATTTTCGGTGTGTTACCTTTTATCCTGTGCTCTGGAAAAGCTATTCGAAGAT
 120
 117 GCATTTTCGGTGTGTTACCTTTTATCCTGTGCTCTGGAAAAGCTATTCGAAGAT
 176
 121 GCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGT
 180
 177 GCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGT
 236
 181 GCTAAAGCAATCATCAACTGCTGTTTATGTATGTAAGCCAGAACAGATCTTATGAGCA
 240
 237 GCTAAAGCAATCATCAACTGCTGTTTATGTATGTAAGCCAGAACAGATCTTATGAGCA
 296
 241 TTGGCACTTCTGTGTGATCTGTTGAGCCGACGACTGAGTGGCTCCAGAACCTTGA
 300
 297 TTGGCACTTCTGTGTGATCTGTTGAGCCGACGACTGAGTGGCTCCAGAACCTTGA
 356
 301 GCCATCCAAATTATGTACCAAAAACCTGCGAGCAAGTGGGCTGAGAAAGTTCACTGGAG
 360
 357 GCCATCCAAATTATGTACCAAAAACCTGCGAGCAAGTGGGCTGAGAAAGTTCACTGGAG
 416
 361 CCACTGAGAAATACCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGACCA
 420
 417 CCACTGAGAAATACCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGACCA
 476
 421 ATTATATAGATAGCATCTCTGGGCTTGGCGACGACATTTGGGACTCTCTCCAGAAAGCA
 480
 477 ATTATATAGATAGCATCTCTGGGCTTGGCGACGACATTTGGGACTCTCTCCAGAAAGCA
 536

```

Qy 481 ACAGCAGAAAGTTCTGTTGTTGACCTTTTCATGATGACAGAGAGGGCTCAGAAACA 540
Db 537 ACAGCAGAAAGTTCTGTTGTTGACCTTTTCATGATGACAGAGAGGGCTCAGAAACA 596
Qy 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
Db 597 AGAGGGAAGATTGTTGTTTATTAACCA 623

RESULT 3
LOCUS AX006440 1860 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 5 from Patient W00004157.
ACCESSION AX006440
VERSION AX006440.1 GI:9994575
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Neefs,J.M., Peeters,D.C. and Pangalos,M.
TITLE Cloning and characterization of novel mammalian peptidases
JOURNAL Patent: WO 0004157-A 5 27-JAN-2000;
JANSSEN PHARMACEUTICA NV (BE); NEEFS JEAN MARC EDMOND FERNAND (BE);
PEETERS DANIELLE CELINE GEORGE (BE); PANGALOS MENLAS (GB)

FEATURES
Source Location/Qualifiers
1..1860
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
471 t

BASE COUNT 518 a 420 c 451 g 471 t
ORIGIN

Query Match 100.0%; Score 567; DB 6; Length 1860;
Best Local Similarity 100.0%; Pred. No. 7.3e-147;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTATCAGATTATCTTAAAGAAAGAAACCACTGGAGAAAAAATGAATTCCTTATCTTC 60
Db 83 CCTATCAGATTATCTTAAAGAAAGAAACCACTGGAGAAAAAATGAATTCCTTATCTTC 142
Qy 61 GCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGAT 120
Db 143 GCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGAT 202
Qy 121 GGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTT 180
Db 203 GGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTT 262
Qy 181 GCTAAGCAATCATCAACTAGCTGTTTATGTTAAAGCCAGAAAGATCTTATGAGGA 240
Db 263 GCTAAGCAATCATCAACTAGCTGTTTATGTTAAAGCCAGAAAGATCTTATGAGGA 322
Qy 241 TTGGACATCTTCTGTTGATGATCTTGAACCAAGCTGATGCTCTCAAGAACTTGAAGAA 300
Db 323 TTGGACATCTTCTGTTGATGATCTTGAACCAAGCTGATGCTCTCAAGAACTTGAAGAA 382
Qy 301 GGCATCCAAATTAATGTAACCAAACTGACAGCAAGATGGCTGAGAAAGTTCACTGGAG 360
Db 383 GGCATCCAAATTAATGTAACCAAACTGACAGCAAGATGGCTGAGAAAGTTCACTGGAG 442
Qy 361 CCAAGTGAATATCCCACTGGAGAGAGGAGAGAGATCAAGCTGTGATCTGAGAGCAAGA 420
Db 443 CCAAGTGAATATCCCACTGGAGAGAGGAGAGAGATCAAGCTGTGATCTGAGAGCAAGA 502
Qy 421 ATTATATAGATAGCATCTGCTGCTTGGAGAGAGCATTTGGAGCTCTCCAGAAAGGATT 480
Db 503 ATTATATAGATAGCATCTGCTGCTTGGAGAGAGCATTTGGAGCTCTCCAGAAAGGATT 562
Qy 481 ACAGCAGAAAGTTCTGTTGTTGACCTTTTCATGATGACAGAGAGGGCTCAGAAACA 540

```

```

Db 563 ACAGCAGAAAGTTCTGTTGTTGACCTTTTCATGATGACAGAGAGGGCTCAGAAACA 622
Qy 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
Db 623 AGAGGGAAGATTGTTGTTTATTAACCA 649

RESULT 4
LOCUS BC020689 1928 bp mRNA linear PRI 22-JAN-2002
DEFINITION Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418
IMAGE:4251802, mRNA, complete cds.
ACCESSION BC020689
VERSION BC020689.1 GI:18088383
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILN)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILN at: http://image.llnl.gov
Series: IRAL Plate: 35 Row: f Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7706386.

FEATURES
Source Location/Qualifiers
1..1928
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/clone="MGC:22418 IMAGE:4251802"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_id="NIH MGC_62"
/lab_host="DH10B"
/notes="vector: pDNR-LIB"
167..1585
/codon_start=1
/product="plasma glutamate carboxypeptidase"
/protein_id="AAH20689.1"
/db_xref="GI:18088384"
/translation="MKFLFAPFGGVHLISGSKAIKRNISKRPFEIKETIASCG
DVAKATINLAVYGAQKQNSYERLALLVDYTPRSLGSKNLEKAIQIMYQVODGLEK
VHLEPVIPIMERGEBSAVLEPFIHITAIIGLSSSITPBGITAEVLVTSFDELQ
RRASEARGKTVVNPQYINYSRTVQQAVERAKVGAALSLRVSASFISYSPHG
IOEVQDGVPIPTACTIVBDVEMMSRASHGIKIVLKNKAGATYPTDSFNTVAET
GSKYDGVVAVSGHLSMDVGGAMDDGGAFISMALSLIKDLGLRPKFTLRVLEMT
ABRGGVAVFOYVQIHRKVNISNYSLVMSAGTFLPTGLPTGSEKRAIIMEEWSLL
QPLNITVYLSHGSEDTINPVIQACVPASLIDLITKFFPHHSIGDTNIMVDPQNAV
AAAVAVASTVVADEMEIPLRS"

BASE COUNT 554 a 435 c 466 g 473 t
ORIGIN

Query Match 100.0%; Score 567; DB 9; Length 1928;

```

Best Local Similarity 100.0%; Pred. No. 7.3e-147;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTATCAGATTATCTTAACAAGAAACCACTGAGAAAAAATGAATTCCTTATCTTC 60
DB 125 CCTATCAGATTATCTTAACAAGAAACCACTGAGAAAAAATGAATTCCTTATCTTC 184
QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAGAACTATATGCAAGAT 120
DB 185 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAGAACTATATGCAAGAT 244
QY 121 GGCATCTCTAAGAGACCTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 180
DB 245 GGCATCTCTAAGAGACCTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 304
QY 181 GCTAAGCAATCATCAACCTAGCTGTTATATGTTAAAGCCAGAAAGATCCATAGAGCA 240
DB 305 GCTAAGCAATCATCAACCTAGCTGTTATATGTTAAAGCCAGAAAGATCCATAGAGCA 364
QY 241 TTGGCACTTCTGTGTGATACCTGTGAGACCAGACTGAGTGGCTCCAGAAACCTAGAAAA 300
DB 365 TTGGCACTTCTGTGTGATACCTGTGAGACCAGACTGAGTGGCTCCAGAAACCTAGAAAA 424
QY 301 GGCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAG 360
DB 425 GGCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAG 484
QY 361 CCAAGTGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTGATGCTGAGACCAAGA 420
DB 485 CCAAGTGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTGATGCTGAGACCAAGA 544
QY 421 ATTCAATGAATAGCATCCTGGCTCTGGACAGACATGGAGCTCCTCCAGAAAGCAATT 480
DB 545 ATTCAATGAATAGCATCCTGGCTCTGGACAGACATGGAGCTCCTCCAGAAAGCAATT 604
QY 481 ACAGCAGAAAGTTCTGTGTGATACCTCTTTCATGAACTGACAGAAAGGCTCAGAAACA 540
DB 605 ACAGCAGAAAGTTCTGTGTGATACCTCTTTCATGAACTGACAGAAAGGCTCAGAAACA 664
QY 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
DB 665 AGAGGGAAGATTGTTGTTTATTAACCA 691

RESULT 5

LOCUS BD127520 1860 bp DNA linear PAT 18-SEP-2002

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD127520.1 GI:23222465

VERSION JP 2002017375-A/2951

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Primer for synthesizing full-length cDNA and use thereof
HEBIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2951
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOUJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC

10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key

FEATURES

source

1. 1860
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FT

CDS

Location/Qualifiers

(129).. (1544).

BASE COUNT 515 a 423 c 451 g 471 t

ORIGIN

Query Match 99.4%; Score 563.8; DB 6; Length 1860;
Best Local Similarity 99.6%; Pred. No. 5.7e-146;
Matches 565; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTATCAGATTATCTTAACAAGAAACCACTGAGAAAAAATGAATTCCTTATCTTC 60
DB 87 CCTATCAGATTATCTTAACAAGAAACCACTGAGAAAAAATGAATTCCTTATCTTC 146
QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAGAACTATATGCAAGAT 120
DB 147 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAGAACTATATGCAAGAT 206
QY 121 GGCATCTCTAAGAGACCTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 180
DB 207 GGCATCTCTAAGAGACCTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 266
QY 181 GCTAAGCAATCATCAACCTAGCTGTTATATGTTAAAGCCAGAAAGATCCATAGAGCA 240
DB 267 GCTAAGCAATCATCAACCTAGCTGTTATATGTTAAAGCCAGAAAGATCCATAGAGCA 326
QY 241 TTGGCACTTCTGTGTGATACCTGTGAGACCAGACTGAGTGGCTCCAGAAAGCAATT 300
DB 327 TTGGCACTTCTGTGTGATACCTGTGAGACCAGACTGAGTGGCTCCAGAAAGCAATT 386
QY 301 GGCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAG 360
DB 387 GGCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAG 446
QY 361 CCAAGTGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTGATGCTGAGACCAAGA 420
DB 447 CCAAGTGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTGATGCTGAGACCAAGA 506
QY 421 ATTCAATGAATAGCATCCTGGCTCTGGACAGACATGGAGCTCCTCCAGAAAGCAATT 480
DB 507 ATTCAATGAATAGCATCCTGGCTCTGGACAGACATGGAGCTCCTCCAGAAAGCAATT 566
QY 481 ACAGCAGAAAGTTCTGTGTGATACCTCTTTCATGAACTGACAGAAAGGCTCAGAAACA 540
DB 567 ACAGCAGAAAGTTCTGTGTGATACCTCTTTCATGAACTGACAGAAAGGCTCAGAAACA 626
QY 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
DB 627 AGAGGGAAGATTGTTGTTTATTAACCA 653

RESULT 6

LOCUS AK075132

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AK075132 1860 bp mRNA linear PRI 03-SEP-2002
Homo sapiens cDNA FLJ90651 fis, clone P14631.004482, moderately
similar to Rattus norvegicus hematopoietic lineage switch 2 related
protein (H182-rp) mRNA.
AK075132
AK075132.1 GI:22761022
oligo capping; f1s (full insert sequence).
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,

Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahari, K., Masuno, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1860)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection; Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1..1860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004482"
/issue_type="placenta"
/clone_id="PLACE1"
/note="Cloning vector: PME18SFL3"
129..1547
/note="unnamed protein product"
/codon_start=1
/protein_id="PAC1423.1"
/db_xref="GI:22761023"
/translation="MKFLIFAFGCVHLLSCSKAI CKNGISKRTFEELIEETASCG
DVAKAIINLVYGAQNSYERLALVDVTPRSGSKNEKAQIMYQNIQDGLER
VHLEPVRI PHMERGESAVMLEPRIHKLAIGLSSIGTPEGITAVLVTSPEDEK
RRASEARGLIVVNOPIYNSRTYQVGAKEAAGALASLRSVSPISYSPHG
IOEVQDQPKIPKACITVEPAEMMSRMSHGIKIVIQKAGKAPPTDSFNVAET
GSKYPEOVVLSHLDMDVQGMADGGGAFISWLSLIKGLPKKTLRLVLT
AEEQGVGARQYQJLHKVNI SNYSLVMSDAGTFLPTGLPTGSEKRAVAEEVMSLL
QPNITQVLSHGEQDINFWIQAGVPSASLDDLKYKFFHSHGDTMTWDPKQNV
AAAVVAVSYVADMEEMLPSS"

BASE COUNT 515 a 423 c 451 g 471 t
ORIGIN

Query Match 99.4%; Score 563.8; DB 9; Length 1860;
Best Local Similarity 99.6%; Pred. No. 5.7e-146;
Matches 565; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60
DB 87 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 146
QY 61 GCATTTTGGTGGTGTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
DB 147 GCATTTTGGTGGTGTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 206
QY 121 GGCATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 180
DB 207 GGCATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 266
QY 181 GCTAAAGCAATCATCACTTACCTGTTTATAGTAAAGCCCAAGACATCTTATGAGCGA 240
DB 267 GCTAAAGCAATCATCACTTACCTGTTTATAGTAAAGCCCAAGACATCTTATGAGCGA 326
QY 241 TTGGCACTTCTGTTGATGTTGAGCCAGACTGAGTGGCTCCAAAGACCTAGAAA 300
DB 327 TTGGCACTTCTGTTGATGTTGAGCCAGACTGAGTGGCTCCAAAGACCTAGAAA 386
QY 301 GCCATCCAAATTATATACAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAG 360
DB 387 GCCATCCAAATTATATACAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAG 446

QY 361 CCAGTGAAGATACCCCACTGGGAGAGGAGAAATCAGCTGTGATGCTGAGCCAGA 420
DB 447 CCAGTGAAGATACCCCACTGGGAGAGGAGAAATCAGCTGTGATGCTGAGCCAGA 506
QY 421 ATTCATTAAGTACCATCTCTGGTCTTGGCAGCAGCATTTGGACCTCTCCAGAGCAT 480
DB 507 ATTCATTAAGTACCATCTCTGGTCTTGGCAGCAGCATTTGGACCTCTCCAGAGCAT 566
QY 481 ACAGCAGAACTTCTGTTGAGACCTCTTTCGATGAACCTGAGAGAGGCTCAGAGCA 540
DB 567 ACAGCAGAACTTCTGTTGAGACCTCTTTCGATGAACCTGAGAGAGGCTCAGAGCA 626
QY 541 AGAGGAGAACTTGTGTTTATACCA 567
DB 627 AGAGGAGAACTTGTGTTTATACCA 653

RESULT 7
BD059610 609 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Secreted expressed sequence tags (ESTs).
ACCESSION BD059610
VERSION BD059610.1 GI:22605216
KEYWORDS JP 2001519666-A/1465.
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 609)
JACOBS, K., MCCOY, J.M., LAVALLIE, E.R., RACIE, L.A., MERBERG, D.,
TREACY, M., SPAULDING, V. and AGOSTINO, M.J.
Secreted expressed sequence tags (ESTs)
Patent: JP 2001519666-A 1465 23-OCT-2001;
GENETICS INSTITUTE INC
PN JP 2001519666-A/1465
PD 23-OCT-2001
PR 10-APR-1998 JP 1998543068
PR 10-APR-1997 US 08/835913
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:
Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..609
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"

BASE COUNT 190 a 127 c 148 g 144 t
ORIGIN

Query Match 97.9%; Score 555; DB 6; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.5e-143;
Matches 566; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60
DB 29 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 88
QY 61 GCATTTTGGTGGTGTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
DB 89 GCATTTTGGTGGTGTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 148
QY 121 GGCATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 180
DB 149 GGCATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 208
QY 181 GCTAAAGCAATCATCACTTACCTGTTTATAGTAAAGCCCAAGACATCTTATGAGCGA 240

Db 209 GCTAAGCAATCATCAACCTAGCTGTTATGTGTAAAGCCCAAGACATCTATGAGCA 268
 QY 241 TTGGACCTTCGTGTGATATCTGTGTGAGACCACTGAGTGGCTCCAGAACTAGAAAA 300
 Db 269 TTGGACCTTCGTGTGATATCTGTGTGAGACCACTGAGTGGCTCCAGAA-CTAGAAAA 327
 QY 301 GCCATCCAAATTTATGTACCAAAAACCTGCAGCAAGTGGCTGAGAAAGTTCACTGGAG 360
 Db 328 GCCATCCAAATTTATGTACCAAAAACCTGCAGCAAGTGGCTGAGAAAGTTCACTGGAG 387
 QY 361 CCAGTGAGAAATCCCCACCTGGAGAGGGGAGAAAGTAATCACTGTGTATGCTGGAGCCAA 420
 Db 388 CCAGTGAGAAATCCCCACCTGGAGAGGGGAGAAAGTAATCACTGTGTATGCTGGAGCCAA 447
 QY 421 ATTCATTAAGATAGCATCTGGGTCTTGGACAGACATGGGACTCTCCAGAGGCATT 480
 Db 448 ATTCATTAAGATAGCATCTGGGTCTTGGACAGACATGGGACTCTCTCCAGAGGCATT 507
 QY 481 ACAGCAGAAATCTGTGTGTGACCTCTTTCATGATGAATGACAGAGAGGCTCAGAAAGCA 540
 Db 508 ACAGCAGAAATCTGTGTGTGACCTCTTTCATGATGAATGACAGAGAGGCTCAGAAAGCA 567
 QY 541 AGAGGAGAAATGTTGTTTATTAACCA 567
 Db 568 AGAGGAGAAATGTTGTTTATTAACCA 594
 RESULT 8
 AF119386 1794 bp mRNA linear PRI 20-MAY-1999
 LOCUS AF119386 Homo sapiens blood plasma glutamate carboxypeptidase precursor
 DEFINITION (PGCP) mRNA, complete cds.
 ACCESSION AF119386
 VERSION AF119386.1 GI:4877697
 KEYWORDS
 ORGANISM Homo sapiens (human)
 SOURCE Homo sapiens
 Bacteria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1794)
 Gindraas, R., Richard, C., El-Alfy, M., Morales, C.R., Potier, M. and Penezhetsky, A.V.
 Purification, cDNA cloning, and expression of a new human blood plasma glutamate carboxypeptidase homologous to N-acetyl-aspartyl-alpha-glutamate carboxypeptidase/prostate-specific membrane antigen
 J. Biol. Chem. 274 (17), 11742-11750 (1999)
 JOURNAL MEDLINE 5923495
 PUBMED 10206990
 2 (bases 1 to 1794)
 Gindraas, R., Richard, C., El-Alfy, M., Morales, C.R., Potier, M. and Penezhetsky, A.V.
 Direct Submission
 Submitted (12-JAN-1999) Medical Genetics, Sainte-Justine Hospital, Montreal University, 3175 Cote Sainte-Catherine, Montreal, QU H3T 1C5, Canada
 TITLE
 JOURNAL
 FEATURES
 source
 1..1794
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="hg22.2; between DBS257 and DBS270"
 1..1794
 /gene="PGCP"
 /gene="PGCP"
 25..1650
 /EC_number="3.4.17.21"
 /function="endopeptidase"
 /note="Member of M28 peptidase family; similar to glutamate carboxypeptidase II and to N-acetyl-aspartyl-alpha-glutamate carboxypeptidase (NALADase); similar to sequences with GenBank Accession

Numbers R18560 and G07404; prostate-specific membrane antigen (PSMA)
 /codon_start=1
 /product="blood plasma glutamate carboxypeptidase precursor"
 /protein_id="AA031418.1"
 /db_xref="GI:4877698"
 /translation="MKFLIPAFPGVHLLSLCSGKAIKNGISKRETEIKIEIASCG DVAKINILAVYGAQRNSYERLALVDVTPGRLSGKNLEKAIQIYVONLODGLK VLEPVRIPMERGESANVLEPRILKILALIGSSITGPEGITAEVLVTSFDEIO RRASERAKIVNQPITYNSTRVQYRTQGAWEAKGALASIRVSASRISYPHTG IOEYQDVVPKIPYACITVEDAEEMSRASHGKIVLQKGAQYPTDSEFNVAET GSKYREQVIVVSGHLDSDMDVGOAMDDGGAFLSMELSLIKDLGAPKSLRLVMT AEOGVGVAPQYQYQHLKVNISYSLVMEASAGTFLPGLQFTGSEKARAIMESVSL OPLNITOVLSHGETDINPFIQAGVPASLLDLYKYPFPHSHGDTMTVMDPSRML LILFGLFIMILQYWKKCCIGPRNSKETSCEPRLGILQIMKTPHITISSSSSK HNSISCFLLISFLILSRKSRKNSHPLEPTT"
 25..96
 /gene="PGCP"
 157..1647
 /gene="PGCP"
 /product="blood plasma glutamate carboxypeptidase"
 205..207
 /gene="PGCP"
 /note="putative; glycosylation site"
 559..561
 /gene="PGCP"
 /note="putative; glycosylation site"
 1081..1083
 /gene="PGCP"
 /note="putative; glycosylation site"
 1090..1092
 /gene="PGCP"
 /note="putative; glycosylation site"
 1210..1212
 /gene="PGCP"
 /note="putative; glycosylation site"
 BASE COUNT 522 a 391 c 422 g 457 t 2 others
 ORIGIN
 Query Match 96.8%; Score 549; DB 9; Length 1794;
 Best Local Similarity 100.0%; Pred. No. 7.5e-142;
 Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 CAAGAAACCACTGGAAGAAAAAATGAATTCCTTATCTCGATTCTTGGTGTGTT 78
 Db 1 CAAGAAACCACTGGAAGAAAAAATGAATTCCTTATCTCGATTCTTGGTGTGTT 60
 QY 79 CACCTTTATCTGTGCTCTGGAGAACTATATGCAAGATGCAATCTCTAAGAGACT 138
 Db 61 CACCTTTATCTGTGCTCTGGAGAACTATATGCAAGATGCAATCTCTAAGAGACT 120
 QY 139 TTGAGAAATTAAGAAATATGCCAGCTGTGGAATGTGCTTAAGCAATATATCAAC 198
 Db 121 TTGAGAAATTAAGAAATATGCCAGCTGTGGAATGTGCTTAAGCAATATATCAAC 180
 QY 199 CTACCTGTTATGTGTAAGCCCAAGACAGATCTATGAGGAGTGGCACTTCTGTTGAT 258
 Db 181 CTACCTGTTATGTGTAAGCCCAAGACAGATCTATGAGGAGTGGCACTTCTGTTGAT 240
 QY 259 ACTGTGGAACCACTGAGTGTGCTCCAGAAAGCTATAGAAAAAGCATCAATATATGATC 318
 Db 241 ACTGTGGAACCACTGAGTGTGCTCCAGAAAGCTATAGAAAAAGCATCAATATATGATC 300
 QY 319 CAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACCTGAGGAGAGTGAATATCCCCAC 378
 Db 301 CAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACCTGAGGAGAGTGAATATCCCCAC 360
 QY 379 TGGAGAGGGAGAAAGATACGTGTATGTGAGAGCAAGATTCATTAAGATGGCATC 438
 Db 361 TGGAGAGGGAGAAAGATACGTGTATGTGAGAGCAAGATTCATTAAGATGGCATC 420
 QY 439 CTGGGTCTTGACACACATTTGGAGCTCTCCAGAAAGCATTAAGAGAAATTTCTGAGT 498

|||||
Db 421 CTGGGCTTGGCAGCAGCATGGGACTCTCCAGAGGCAATTAACAGCAAGATTCTGGTG 480
Qy 499 GTACACCTCTTCCATGATGACGAGAAAGGGCCCTCAGAAAGCAAGGAATTTGTTG 558
Db 481 GTGACCTCTTTCATGATGACGAGAAAGGGCCCTCAGAAAGCAAGGAATTTGTTG 540
Qy 559 TATACCAA 567
Db 541 TATACCAA 549

RESULT 9
BD105816
LOCUS
DEFINITION Secretory protein.
ACCESSION BD105816
VERSION BD105816.1 GI:23200634
KEYWORDS JP 2002502234-A/7.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1778)
AUTHORS Jacobs,K., McCoy,J.M., Racie,L.A., Lavallie,E.R., Merberg,D. and Spaulding,V.
TITLE Secretory protein
JOURNAL Patent: JP 2002502234-A 7 22-JAN-2002;
COMMENT GENETICS INSTITUTE INC
PN JP 2002502234-A/7
PD 22-JAN-2002
PF 16-APR-1997 JP 1997537384
PR 18-APR-1996 US 08/634325 13-JAN-1997 US 08/783520 PI
KENNETH JACOBS,JOHN M MCCOY,LISA A RACIE,EDWARD R LAVALLIE, PI
DAVID MERBERG,
PI VIKKI SPAULDING
PC C12N15/12,C07K14/47,A61K38/17
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
SOURCE 1..1778
/organism="synthetic construct"
/mol type="genomic DNA"
/db xref="taxon:32630"
BASE COUNT 514 a 386 c 422 g 456 t
ORIGIN
Query Match 92.8%; Score 526.2; DB 6; Length 1778;
Best Local Similarity 98.5%; Pred. No. 1.7e-135;
Matches 531; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 29 AACTGGAAGAAAAAATGAAATTCCTTATCTTGCGATTTTGGTGCTGTCACCTTTAT 88
Db 3 AGCTTGGGACAGAGATGAAATTCCTTATCTTGCGATTTTGGTGCTGTCACCTTTAT 62
Qy 89 CCCGTGCTGGGAGAACTATATGCAAGATGCGATCTTAAGGCACTTTGAAGAA 148
Db 63 CCTGTGCTGGGAGAACTATATGCAAGATGCGATCTTAAGGCACTTTGAAGAA 122
Qy 149 TAAAGAAGAAATAGCCAGCTGTGGAGATGTGCTAAAGCAATCAACCTAGCTGTT 208
Db 123 TAAAGAAGAAATAGCCAGCTGTGGAGATGTGCTAAAGCAATCAACCTAGCTGTT 182
Qy 209 ATGTAAAGCCCAAGACAGATCTATAGCAATTTGGCACTTCTGTGATATCTGTTGAC 268
Db 183 ATGTAAAGCCCAAGACAGATCTATAGCAATTTGGCACTTCTGTGATATCTGTTGAC 242
Qy 269 CCAGCTAGTGGCTCCCAAGAACCTTAAGAAAGCAATCAAAATTAATGACAAAACCTGC 328
Db 243 CCAGCTAGTGGCTCCCAAGAACCTTAAGAAAGCAATCAAAATTAATGACAAAACCTGC 302
Qy 329 AGCAAGATGGGCTGAGAAAGTTCACTGAGCCAGTGAAGATACCCCACTGGAGAGGG 388

|||||
Db 303 AGCAAGATGGGCTGAGAAAGTTCACTGAGCCAGTGAAGATACCCCACTGGAGAGGG 362
Qy 389 GAGAGAAATCAGCTGTATGATGAGCCAGAAATTCATAGATTAAGCATCTGGGCTTG 448
Db 363 GAGAGAAATCAGCTGTATGATGAGCCAGAAATTCATAGATTAAGCATCTGGGCTTG 422
Qy 449 GCAGCAGCAATTTGGACTCTCCAGAAAGCAATTAACAGCAAGATTTGGTGACCTTT 508
Db 423 GCAGCAGCAATTTGGACTCTCCAGAAAGCAATTAACAGCAAGATTTGGTGACCTTT 482
Qy 509 TCGATGAATCTGCAGAGAAAGGCTCCAGAGCAAGAGGAAGATTGTTTATACCAA 567
Db 483 TCGATGAATCTGCAGAGAAAGGCTCCAGAGCAAGAGGAAGATTGTTTATACCAA 541

RESULT 10
BD125210
LOCUS
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD125210
VERSION BD125210.1 GI:23220155
KEYWORDS JP 2002017375-A/641.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 642)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 641 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/641
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOgai,KOJI HAYASHI,SHIZUKO PI
ISHII,
PI YURI KAWAI,AI MAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10,
C12P21/02,C12P21/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
SOURCE 1..642
/organism="Homo sapiens"
/mol type="genomic DNA"
/db xref="taxon:9606"
BASE COUNT 183 a 144 c 171 g 139 t 5 others
ORIGIN
Query Match 90.3%; Score 512.2; DB 6; Length 642;
Best Local Similarity 98.4%; Pred. No. 1.2e-131;
Matches 547; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

Qy 1 CCTATCAGATTAATCTTAACAAGAAACCACTGGAAGAAAGAAATTCCTTATCTTC 60
Db 87 CCTATCAGATTAATCTTAACAAGAAACCACTGGAAGAAAGAAATTCCTTATCTTC 146
Qy 61 GATTTTGGTGAGTGTACCTTTTATCCCTGTGCTCTGGGAAGCAATTAAGCAAGAT 120
Db 147 GATTTTGGTGAGTGTGTACCTTTTATCCCTGTGCTCTGGGAAGCAATTAAGCAAGAT 206
Qy 121 GGCAATCTTAAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGAAGATTT 180
Db 207 GGCAATCTTAAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGAAGATTT 266

Query Match 90.3%; Score 512.2; DB 6; Length 642;
 Best Local Similarity 98.4%; Pred. No. 1.2e-11;
 Matches 547; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 181 GCTAAGCATCATCAACCTAGCTGTTATGTAAAGCCAGAACAGATCCTATGAGCA 240
 DB 267 GCTAAGCATCATCAACCTAGCTGTTATGTAAAGCCAGAACAGATCCTATGAGCA 326
 QY 241 TTGGCAGCTTCTGTTGATGATGAGCCAGACTGATGCTCCAGAACCTAGAAAA 300
 DB 327 TTGGCAGCTTCTGTTGATGATGAGCCAGACTGATGCTCCAGAACCTAGAAAA 386
 QY 301 GCCATCCAAATTATGTACCAAACTCGACGAGATGGGCTGGAGAAAAGTTCACTGAG 360
 DB 387 GCCATCCAAATTATGTACCAAACTCGACGAGATGGGCTGGAGAAAAGTTCACTGAG 446
 QY 361 CCAGTGAATATACCCACTGGGAGAGGGAGAGAAATAGCTGTATGCTGAGCAAGA 420
 DB 447 CCAGTGAATATACCCACTGGGAGAGGGAGAGAAATAGCTGTATGCTGAGCAAGA 506
 QY 421 ATTGATTA-GATAGCATCTGCTGCTTGGAGCAGCATTTGGGACTCCTCCAGAGGAT 479
 DB 507 ATTGATTA-GATAGCATCTGCTGCTTGGAGCAGCATTTGGGACTCCTCCAGAGGAT 566
 QY 480 TACAGCAGAGTTCTGTGTG-TGACCTCTTTCGATGAAGTGCAGAGAA-GGGCTCAGAA 537
 DB 567 TACAGCAGAGTTCTGTGTGTG-TGACCTCTTTCGATGAAGTGCAGAGAAAGGGCTCAGAA 626
 QY 538 GCAGAGGGAGAGATTG 553
 DB 627 GCAGAGGGAGAGATTG 642

RESULT 11
 BD126476 642 bp DNA linear PAT 18-SEP-2002
 LOCUS BD126476
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD126476
 VERSION 1 GI:232221421
 KEYWORDS JP 2002017375-A/1907.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
 Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
 Koga,H.
 TITLE Primer for synthesizing full-length cDNA and use thereof
 JOURNAL Patent: JP 2002017375-A 1907 22-JAN-2002.
 COMMENT HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002017375-A/1907
 PD 22-JAN-2002 JP 2002051172
 PF 07-JUL-2000 JP 2002051172
 PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
 PI ISHII,
 PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
 SHINICHI KOJIMA,
 PI TETSUO OTSUKI, HISASHI KOGA

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/00, C12N15/00, C06F17/30, C12N15/00, C12N5/00 CC
 PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 FT source 1. 642
 FT Location/Qualifiers
 /organism="Homo sapiens (human)"

FEATURES
 source 1. 642
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 183 a 144 c 171 g 139 t 5 others
 ORIGIN

Query Match 90.3%; Score 512.2; DB 6; Length 642;
 Best Local Similarity 98.4%; Pred. No. 1.2e-11;
 Matches 547; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 1 CCTATCAGATTATCTTAAAGAAACCACTGAAAAAATGAAATTCCTTATCTTC 60
 DB 87 CTTATCAGATTATCTTAAAGAAACCACTGAAAAAATGAAATTCCTTATCTTC 146
 QY 61 GCATTTTCGGTGTGTCACCTTTTATCCCTGTGCTGGGAAAGCTATATGCAAGAT 120
 DB 147 GCATTTTCGGTGTGTCACCTTTTATCCCTGTGCTGGGAAAGCTATATGCAAGAT 206
 QY 121 GGCATCTCTAAGAGACTTTTGAAGAAATTAAGAGAAATAGCCACTGTGAGATGT 180
 DB 207 GGCATCTCTAAGAGACTTTTGAAGAAATTAAGAGAAATAGCCACTGTGAGATGT 266
 QY 181 GCTAAGCATCATCAACCTAGCTGTTATGTAAAGCCAGAACAGATCCTATGAGCA 240
 DB 267 GCTAAGCATCATCAACCTAGCTGTTATGTAAAGCCAGAACAGATCCTATGAGCA 326
 QY 241 TTGGCAGCTTCTGTTGATGATGAGCCAGACTGATGCTCCAGAACCTAGAAAA 300
 DB 327 TTGGCAGCTTCTGTTGATGATGAGCCAGACTGATGCTCCAGAACCTAGAAAA 386
 QY 301 GCCATCCAAATTATGTACCAAACTCGACGAGATGGGCTGGAGAAAAGTTCACTGAG 360
 DB 387 GCCATCCAAATTATGTACCAAACTCGACGAGATGGGCTGGAGAAAAGTTCACTGAG 446
 QY 361 CCAGTGAATATACCCACTGGGAGAGGGAGAGAAATAGCTGTATGCTGAGCAAGA 420
 DB 447 CCAGTGAATATACCCACTGGGAGAGGGAGAGAAATAGCTGTATGCTGAGCAAGA 506
 QY 421 ATTGATTA-GATAGCATCTGCTGCTTGGAGCAGCATTTGGGACTCCTCCAGAGGAT 479
 DB 507 ATTGATTA-GATAGCATCTGCTGCTTGGAGCAGCATTTGGGACTCCTCCAGAGGAT 566
 QY 480 TACAGCAGAGTTCTGTGTG-TGACCTCTTTCGATGAAGTGCAGAGAA-GGGCTCAGAA 537
 DB 567 TACAGCAGAGTTCTGTGTGTG-TGACCTCTTTCGATGAAGTGCAGAGAAAGGGCTCAGAA 626
 QY 538 GCAGAGGGAGAGATTG 553
 DB 627 GCAGAGGGAGAGATTG 642

RESULT 12
 AC084223/c 156702 bp DNA linear HTG 07-JUL-2001
 LOCUS AC084223/c
 DEFINITION Homo sapiens chromosome 8 clone RP11-80P10 map 8, WORKING DRAFT
 SEQUENCE, 4 unordered pieces.
 ACCESSION AC084223
 VERSION 3 GI:14626321
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 TITLE Homo sapiens chromosome 8, clone RP11-80P10
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 156702)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Barina,N., Baetien,V., Beda,F., Boguslavsky,L.,
 Bonkagiller,B., Brown,A., Burkett,G., Campopiano,A., Casale,A.,
 Chepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
 Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
 Graham,L., Grand-pierre,N., Hagos,B., Heathford,A., Horton,L.,
 Iliev,I., Johnson,R., Jones,C., Kam,L., Karas,A., Labrecque,K.,
 Lamazares,R., Landers,T., Lehoczy,D., Levine,R., Liu,C., Liu,G.,
 Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
 McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,

Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivar, T.M., Oliver, J., Peterson, R., Pierre, N., Pisan, C., Pollard, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougnuez, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strassman, N., Subramanian, A., Talamas, J., Teffaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Vei, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

Submitted (17-OCT-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jul 7, 2001 this sequence version replaced gi:13249444. All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: U11261

Center clone name: 80_P_10

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 154179 bases at least Q40

Consensus quality: 155367 bases at least Q30

Consensus quality: 155704 bases at least Q20

Insert size: 163000; agarose-fp

Insert size: 156402; sum-of-ctg

Quality coverage: 5.7 in Q20 bases; agarose-fp

Quality coverage: 6.0 in Q20 bases; sum-of-ctgs

NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 12817: contig of 12817 bp in length
* 12818 12917: gap of 100 bp
* 12918 59770: contig of 46853 bp in length
* 59771 59870: gap of 100 bp
* 59871 113148: contig of 53278 bp in length
* 113149 113248: gap of 100 bp
* 113249 156702: contig of 43454 bp in length.

Location/Qualifiers

1. 156702

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-80P10"

/clone_lib="RPC1-11 Human Male BAC"

1. 12817

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

12918. 59770

/note="assembly_fragment"

59871. 113148

/note="assembly_fragment"

113249. 156702

/note="assembly_fragment"

clone_end:T7

vector_side:right"

BASE COUNT 52002 a 30386 c 29018 g 44993 t 303 others

ORIGIN

Query Match 83.3%; Score 472.4; DB 2; Length 156702;
Best Local Similarity 99.8%; Pred. No. 2,3e-120;
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

7 AGATTATCTTACAGAAAACCACTGGAAAAAATGAATTCCTTATCTTGCATTT 66
63583 AGATTATCTTACAGAAAACCACTGGAAAAAATGAATTCCTTATCTTGCATTT 63524

67 TTGGTGTGTGTTCACTTTATCTTGTCTGTCTGGAAAGCTATATGCAAGATGGCATC 126
63523 TTGGTGTGTGTTCACTTTATCTTGTCTGTCTGGAAAGCTATATGCAAGATGGCATC 63464

127 TCTAGAGAGCTTTGAGAAATTAAGAAATATAGCAGCTGTGGAGATGGTCTTAA 186
63463 TCTAGAGAGCTTTGAGAAATTAAGAAATATAGCAGCTGTGGAGATGGTCTTAA 63404

187 GCATCATCAACCTAGCTTTTATGTGTAAGCCAGAACAGATCTATGAGCCATTGGCA 246
63403 GCATCATCAACCTAGCTTTTATGTGTAAGCCAGAACAGATCTATGAGCCATTGGCA 63344

247 CTCTGTGTGATGATCTGTGGAACCACTAGTGTCTTCAAGAACTTGAAGAAACCATC 306
63343 CTCTGTGTGATGATCTGTGGAACCACTAGTGTCTTCAAGAACTTGAAGAAACCATC 63284

307 CAATTATGTACCAAACTGCAGCAAGATGGGTGGAGAAAGTTCACTGGAGCCAGTG 366
63283 CAATTATGTACCAAACTGCAGCAAGATGGGTGGAGAAAGTTCACTGGAGCCAGTG 63224

367 AGAATACCCAGCTGGAGAGAGGAGAAATACAGCTGTGATGTGAGCCAGAAATTCAT 426
63223 AGAATACCCAGCTGGAGAGAGGAGAAATACAGCTGTGATGTGAGCCAGAAATTCAT 63164

427 AAGATGACCATCTGCTGTCTTGGCAGCAGCATTTGGACTCTCTCAAGAGCATTT 480
63163 AAGATGACCATCTGCTGTCTTGGCAGCAGCATTTGGACTCTCTCAAGAGCATTT 63110

RESULT 13

AP006278

LOCUS

Homo sapiens genomic DNA, chromosome 8q22.1, clone: KB1573B4,

DEFINITION

complete sequence.

ACCESSION

AP006278

AP006278.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1

Shimizu, N. and Asakawa, S.

Homo sapiens DNA chromosome 8 SEQUENCE

PUBLISHED ONLY in Database (2003)

2 (bases 1 to 157927)

Shimizu, N. and Asakawa, S.

Direct Submission

Submitted (24-MAR-2003) Nobuyoshi Shimizu, Keio University, School

of Medicine, Molecular Biology/ 35 Shinanomachi, Shinjuku-ku, Tokyo

160-8582, Japan (E-mail: nhimizu@med.keio.ac.jp,

Tel: 81-3-3351-2370, Fax: 81-3-3351-2370)

Location/Qualifiers

1. 157927

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="8"

/map="8q22.1"

/clone="KB1573B4"

/cell_line="FLB 14 - 14"

/cell_type="pre-pro-B cell"

/clone_lib="Keio BAC library"

repeat_region	103. .202	/evidence=not experimental
repeat_region	/rpt_family="L1ME"	/rpt_family="T1A)n"
repeat_region	complement(217. .695)	/evidence=not experimental
repeat_region	/rpt_family="HERVL"	/rpt_family="AT_rich"
repeat_region	complement(733. .1197)	/evidence=not experimental
repeat_region	/rpt_family="HERVL"	/rpt_family="L1PA11"
repeat_region	complement(1135. .1381)	/evidence=not experimental
repeat_region	/rpt_family="HERVL"	/rpt_family="L1M4"
repeat_region	complement(1462. .1817)	/evidence=not experimental
repeat_region	/rpt_family="L1M4"	/evidence=not experimental
repeat_region	complement(1821. .1863)	/evidence=not experimental
repeat_region	/rpt_family="MLT2A1"	/rpt_family="L1M4"
repeat_region	1864. .1888	/evidence=not experimental
repeat_region	/rpt_family="T1CA)n"	/rpt_family="HER103"
repeat_region	1905. .1961	/evidence=not experimental
repeat_region	/rpt_family="T1A)n"	/rpt_family="MIR"
repeat_region	complement(1973. .2381)	/evidence=not experimental
repeat_region	/rpt_family="MLT2A1"	/rpt_family="L1M4"
repeat_region	complement(2473. .3057)	/evidence=not experimental
repeat_region	/rpt_family="L1M4"	/rpt_family="AT_rich"
repeat_region	3064. .3092	/evidence=not experimental
repeat_region	/rpt_family="AT_rich"	/rpt_family="T1PA3"
repeat_region	complement(3093. .9136)	/evidence=not experimental
repeat_region	/rpt_family="L1PA3"	complement(9154. .9483)
repeat_region	/evidence=not experimental	/rpt_family="L1M4"
repeat_region	9661. .9951	/evidence=not experimental
repeat_region	/rpt_family="ALUSg"	/rpt_family="MIR"
repeat_region	9960. .10103	/evidence=not experimental
repeat_region	/rpt_family="L1G"	/rpt_family="T1A)n"
repeat_region	11642. .11941	/evidence=not experimental
repeat_region	/rpt_family="ALUSg"	/rpt_family="MIR"
repeat_region	12296. .12351	/evidence=not experimental
repeat_region	/rpt_family="MIR"	/rpt_family="L1M4"
repeat_region	12352. .12530	/evidence=not experimental
repeat_region	/rpt_family="MIR3"	/rpt_family="T1A)n"
repeat_region	12531. .12575	/evidence=not experimental
repeat_region	/rpt_family="MIR3"	/rpt_family="L1M4"
repeat_region	13183. .13204	/evidence=not experimental
repeat_region	/rpt_family="AT_rich"	/rpt_family="T1A)n"
repeat_region	13487. .13604	/evidence=not experimental
repeat_region	/rpt_family="L1M4"	/rpt_family="MIR"
repeat_region	14340. .14545	/evidence=not experimental
repeat_region	/rpt_family="L1M4"	/rpt_family="T1A)n"
repeat_region	14753. .14795	/evidence=not experimental
repeat_region	14981. .14981	/evidence=not experimental
repeat_region	/rpt_family="AT_rich"	/rpt_family="L1PA11"
repeat_region	complement(15252. .17419)	/evidence=not experimental
repeat_region	/rpt_family="L1PA11"	/rpt_family="L1M4"
repeat_region	complement(17577. .17931)	/evidence=not experimental
repeat_region	/rpt_family="L1M4"	/rpt_family="T1A)n"
repeat_region	complement(18038. .18415)	/evidence=not experimental
repeat_region	/rpt_family="L1M4"	/rpt_family="L1M4"
repeat_region	complement(18566. .18800)	/evidence=not experimental
repeat_region	/rpt_family="L1M4"	/rpt_family="MIR"
repeat_region	18854. .18968	/evidence=not experimental
repeat_region	/rpt_family="MIR103"	/rpt_family="L1M4"
repeat_region	19398. .19468	/evidence=not experimental
repeat_region	/rpt_family="MIR"	/rpt_family="T1A)n"
repeat_region	complement(19949. .20245)	/evidence=not experimental
repeat_region	/rpt_family="L1M4"	/rpt_family="AT_rich"
repeat_region	20246. .20316	/evidence=not experimental
repeat_region	/rpt_family="AT_rich"	/rpt_family="L1M4"
repeat_region	20503. .20523	/evidence=not experimental
repeat_region	/rpt_family="AT_rich"	/rpt_family="T1A)n"
repeat_region	complement(20530. .20781)	/evidence=not experimental
repeat_region	/rpt_family="HAL1"	/rpt_family="L1M4"
repeat_region	21107. .21135	/evidence=not experimental
repeat_region	/rpt_family="AT_rich"	/rpt_family="MIR"
repeat_region	complement(21428. .21700)	/evidence=not experimental
repeat_region	/rpt_family="HAL1"	/rpt_family="L1M4"
repeat_region	21701. .22034	/evidence=not experimental
repeat_region	/rpt_family="ALUSg"	/rpt_family="T1A)n"
repeat_region	complement(22035. .22140)	/evidence=not experimental
repeat_region	/rpt_family="HAL1"	/rpt_family="L1M4"
repeat_region	complement(24217. .24393)	/evidence=not experimental
repeat_region	/rpt_family="MIR"	/rpt_family="T1A)n"
repeat_region	complement(24453. .24675)	/evidence=not experimental
repeat_region	/rpt_family="MIR"	/rpt_family="L1M4"
repeat_region	complement(24890. .25101)	/evidence=not experimental
repeat_region	/rpt_family="MIR"	/rpt_family="T1A)n"
repeat_region	26204. .26409	/evidence=not experimental
repeat_region	/rpt_family="MIR"	/rpt_family="L1M4"
repeat_region	26451. .26507	/evidence=not experimental
repeat_region	/rpt_family="MIR"	/rpt_family="T1A)n"
repeat_region	26979. .27110	/evidence=not experimental
repeat_region	/rpt_family="MIR"	/rpt_family="L1M4"
repeat_region	complement(27349. .27525)	/evidence=not experimental
repeat_region	/rpt_family="L1M4"	/rpt_family="T1A)n"
repeat_region	complement(27526. .27908)	/evidence=not experimental
repeat_region	/rpt_family="MIR"	/rpt_family="L1M4"
repeat_region	complement(27909. .28143)	/evidence=not experimental


```

|||||
Db      93526 AGAATACCCCACTGGGAGAGGGGAGAAATCAGCTGTGATGCTGGAGCCAGAAATTCAT 93585
QY      427  AAGATGCCATCCCTGGGCTCTTGGCAGACGATTTGGGACTCTCCAGAAAGGATT 480
Db      93586 AAGATGCCATCCCTGGGCTCTTGGCAGACGATTTGGGACTCTCCAGAAAGGATT 93639

RESULT 15
AF131077
LOCUS   AF131077 1726 bp mRNA linear ROD 03-JAN-2003
DEFINITION Rattus norvegicus liver annexin-1-like protein (LAL) mRNA, complete
c88.
ACCESSION AF131077 GI:7108712
VERSION   AF131077.1
KEYWORDS  Rattus norvegicus (Norway rat)
SOURCE    Rattus norvegicus
ORGANISM  Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 1726)
          Della Fazio, M.A., Piobbico, D., Bartoli, D., Castelli, M.,
          Brancorsini, S., Viola Magni, M. and Servillo, G.
          1a1-1: a differentially expressed novel gene during proliferation
          in liver regeneration and in hepatoma cells
          Genes Cells 7 (11), 1183-1190 (2002)
JOURNAL   22278398
MEDLINE   12390252
FUEMED    2 (bases 1 to 1726)
REFERENCE 2 (bases 1 to 1726)
          Servillo, G., Della Fazio, M.A., Piobbico, D., Bartoli, D.,
          Castelli, M., Brancorsini, S. and Viola Magni, M.
          Direct Submission
          Submitted (25-FEB-1999) Institute of General Pathology, University
          of Perugia, Policlinico Monteluce, Perugia 06100, Italy
          Location/Qualifiers
            1..1726
              /organism="Rattus norvegicus"
              /mol_type="mRNA"
              /db_xref="taxon:10116"
            1..1726
              /gene="LAL"
              /gene="LAL"
              /gene="LAL"
              /note="Involved in liver regeneration; contains unique
              annexin domain at the N-terminus"
              /codon_start=1
              /product="liver annexin-1-like protein"
              /protein_id="AAF36518.1"
              /db_xref="GI:7108713"
              /translation="MRPLFLFVAVVHLFSLGSKATYKSGVSORTFOEIKERIANYE
              DVAKATINLAVTKYONRSTYERGLVDIVGPRLSGSKNLEKAIQIMTQNDGLN
              VHLQVRI PHMERGESAVVWVPRIHKLALIGLGSIGTPEGITAEVLVAVSFELQ
              RRASEARGKIVVNPYTDYGVQYRERGAFAKAVASLIRVASFSISVPHTG
              HOGYODGVPKIPACTIIEDAEMMSRTASRGDKIVILKMGAKTYPTDSFNTVAEIT
              GSKYPERVIVSGHLSMDWDGQALDDGGAFLSMELSLKDLGLRPKRTLRVLMT
              AEBGCGVGAASYREILHKANISKISLVWEADSGTFLPGLQPTGSDKRAAIKEVMSLI
              QPLNITVFNDABSTIDINFWIQAGVPGASLKDILYKTFPFHSHGDTMTADPKOMNV
              AAAYVAVVAVVADMEEMLPERS"
BASE COUNT 474 a 385 c 428 g 439 t
ORIGIN
Query Match 70.5%; Score 400; DB 10; Length 1726;
Best Local Similarity 82.8%; Pred. No. 2.4e-100;
Matches 457; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```

```

QY      135  GACTTTGAAGAAATAAAGAAAGAAATAGCCACGTGTGAGATGTTGCTAAAGCAATCAT 194
Db      155  AACTTTCAAGAAATAAAGAAAGAAATAGCCACCTATGAAGATGTTGCTAAAGCAATCAT 214
QY      195  CAACCTAGCTGTTTATGTATAGTAAAGCCAGAAACAGATCCTATGACGATGGCACTTGTGT 254
Db      215  CAACCTAGCTGTTTATGTATAGTAAATATACCAAGACCGGTGATGAGCGTTTGGCACTTGTAGT 274
QY      255  TGATACTGTTGAGACCCAGACTGATGTGGCTCCAGAAAGCTTGAAGAAAGCCATCCAAATAT 314
Db      275  TGATACTGTTGAGACCCAGACTGATGTGGCTCTTAAGAACTTAGAAGAAAGCTTCCAAATCAT 334
QY      315  GTACCAAAAACCTGCAGCAAGATGAGCTGGAGAAAGTTCACTGAGCCAGTGAAGATATACC 374
Db      335  GTACCAAAAACCTGCACACAGATGAGCTGGAGAAAGCTTCACTGAGCCAGTGAAGATATACC 394
QY      375  CCACTGGAGAGAGGAGAGAAATCAAGCTGTGATGCTGGAGCCAGAAATTCATAAGATAGC 434
Db      395  TCACTGGAGAGAGGAGAGAAATCTGCAATGATGTGTGCTCTGAAATTCACAAAGTTGGC 454
QY      435  CATTCTGGCTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCATTAACAGAAAGTTCT 494
Db      455  TATTTTAGGCTTGGCGGAGCATTTGGGACTCTCTGAAAGGTATACAGCAGAAAGTACT 514
QY      495  GGTGTGACCTCTTTCATGATGACCTGCAGAGAGGCTCAGAAAGAGAGGAAATTTGT 554
Db      515  GGTGTGACCTCTTTCATGATGACCTGCAGAGAGGCTCAGAAAGAGGAAATTTGT 574
QY      555  TGTATTATACCA 566
Db      575  TGTATTATACCA 586

```

Search completed: December 22, 2003, 20:22:27
 Job time : 2052.56 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:18:08 ; Search time 6681.9 Seconds
(without alignments)
11332.658 Million cell updates/sec

Title: US-09-745-763-35

Perfect score: 1851

Sequence: 1 GGCTAGCGCCGAGCTTACT.....CTGAAAAAAAAAAAAAAAAA 1851

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*
41: gb_ov:*

```

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1851	100.0	1851	6	BD106411
2	1839.8	99.4	1928	6	BC020689 Homo sapi
3	1820.4	98.3	1860	6	BD127520 Rattus no
4	1820.4	98.3	1860	6	BD127520 Rattus no
5	1783.2	96.3	1860	6	AX006440 Sequence
6	1758.6	95.0	1794	6	AF119386 Homo sapi
7	1751	94.6	1778	6	BD105816 Rattus no
8	1465.6	79.2	1472	6	AF107834 Homo sapi
9	1145	62.9	1796	10	BC037067 Mus muscu
10	1145	61.9	1778	6	AX400829 Sequence
11	1145	61.9	1778	10	AF097723 Rattus no
12	1143.8	61.8	1726	10	AF131077 Rattus no
13	1118.4	60.4	1596	10	AF107835 Mus muscu
14	1077.8	58.2	1716	10	AF009513 Mus muscu
15	569	30.7	609	6	BD059610 Secretd
16	568.2	30.7	642	6	BD125210 Primer fo
17	568.2	30.7	642	6	BD125210 Primer fo
18	484.4	26.2	152176	9	AP003117 Homo sapi
19	484.4	26.2	166050	2	AC013817 Homo sapi
20	484.4	26.2	177008	2	AC011175 Homo sapi
21	472.4	25.5	156702	2	AC084223 Homo sapi
22	472.4	25.5	157927	9	AP006278 Homo sapi
23	472.4	25.5	171301	9	AC010859 Homo sapi
24	472.4	25.5	171301	9	AF107833 Homo sapi
25	410.6	22.2	492	6	BD125961 Primer fo
26	331	17.9	417	10	AF107832 Mus muscu
27	327.4	17.7	167883	2	AC121026 Rattus no
28	327.4	17.7	218874	2	AC110103 Rattus no
29	324.4	17.5	360	6	BD105812 Secretd
30	324.4	17.5	360	6	E62850 Secretd
31	323.6	17.5	214287	2	AC129951 Mus muscu
32	310.2	16.8	317	6	BD076934 5' EST of
33	290.8	15.7	314	6	BD058492 Secretd
34	284.6	15.4	462	6	BD105811 Secretd
35	284.6	15.4	462	6	E62849 Secretd
36	233	12.6	234	6	AX411237 Sequence
37	233	12.6	234	11	G07404 human SRS W
38	212.2	11.5	144703	9	AP003112 Homo sapi
39	205	11.1	171035	9	AP003111 Homo sapi
40	174.4	9.4	214287	2	AC129951 Mus muscu
41	174.4	9.4	214287	2	AC129951 Mus muscu
42	172.8	9.3	203469	2	AC123654 Mus muscu
43	171.2	8.9	218874	2	AC110103 Rattus no
44	163.4	8.8	288136	2	AC120489 Rattus no
45	154.4	8.3	238034	2	AC123681 Mus muscu

ALIGNMENTS

```

RESULT 1
LOCUS      BD106411      1851 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION  BD106411
VERSION    BD106411.1  GI:23201229
KEYWORDS   Chlamydia sp.
SOURCE     Chlamydia sp.
ORGANISM   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE  1 (bases 1 to 1851)
AUTHORS    Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M., Spaulding,V. and Agostino,M.J.
TITLE      Secreted proteins and polynucleotides encoding them
JOURNAL    Patent: JP 2002503955-A 2 05-FEB-2002;

```

Pred. No. is the number of results predicted by chance to have a

GENETICS INSTITUTE INC
PN JP 2002503955-A/2

PD 05-FEB-2002
PF 20-MAR-1998 JP 1998545874
PR 21-MAR-1997 US 08/822167, 19-MAR-1998 US 09/044466 PI
KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACY, VIKKI SPALDING, MICHAEL J AGOSTINO PC
C12N15/12, C07K14/47, A61K38/17
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.

FEATURES
Source 1. 1851
/organism="Chlamydia sp."
/mol_type="genomic DNA"
/db_xref="taxon:35827"

BASE COUNT 531 a 413 c 438 g 469 t
ORIGIN

Query Match 100.0%; Score 1851; DB 6; Length 1851;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCTAGGCGGAGCTTGTCTCTGGAGCCGCTCCGTCGCGCCGTCAGAGCCGCCCTTA 60
1 GGCTAGGCGGAGCTTGTCTCTGGAGCCGCTCCGTCGCGCCGTCAGAGCCGCCCTTA 60
61 TCAGATTAATCTTAACAAGAAAAACAATGGAATAATGAAATTCCTTAATCTTCGCAT 120
61 TCAGATTAATCTTAACAAGAAAAACAATGGAATAATGAAATTCCTTAATCTTCGCAT 120
121 TTTTCGGGGGTTCACTTTTATCCCTGCTGCTGGAGAACTATATGCAAGATGGCA 180
121 TTTTCGGGGGTTCACTTTTATCCCTGCTGCTGGAGAACTATATGCAAGATGGCA 180
121 TTTTCGGGGGTTCACTTTTATCCCTGCTGCTGGAGAACTATATGCAAGATGGCA 180
181 TCTCTAAGAGAGCTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGTTGCTA 240
181 TCTCTAAGAGAGCTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGTTGCTA 240
181 TCTCTAAGAGAGCTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGTTGCTA 240
241 AAGCAATCATCAACCTGATCTGTTATGTTAAAGCCCAAGACAGATCTATGAGCATTTGG 300
241 AAGCAATCATCAACCTGATCTGTTATGTTAAAGCCCAAGACAGATCTATGAGCATTTGG 300
301 GACTTCGTGGTGAATCTGTGGACCACTGAGTGGCTCCAAAGAACTTAAGAAAGCA 360
301 GACTTCGTGGTGAATCTGTGGACCACTGAGTGGCTCCAAAGAACTTAAGAAAGCA 360
361 TCCAAATATGATGACCAAAACCTGACAGCAAGATGGCTGGAAGAAATGCACTGAGCCAG 420
361 TCCAAATATGATGACCAAAACCTGACAGCAAGATGGCTGGAAGAAATGCACTGAGCCAG 420
421 TGAAGATATCCCACTGGAGAGGAGGAAGAAATCACTGTGATGCTGAGCCAAAGATTC 480
421 TGAAGATATCCCACTGGAGAGGAGGAAGAAATCACTGTGATGCTGAGCCAAAGATTC 480
481 ATAAGATAGCCATCTGGGCTCTTGGCAGCAGATTTGGACCTCTCCAGAAAGCATTAAG 540
481 ATAAGATAGCCATCTGGGCTCTTGGCAGCAGATTTGGACCTCTCCAGAAAGCATTAAG 540
541 CAGAAGTCTGGTGGTGAATCTTTCGATGAACTCAAGAGAGGAGCTCCAGAGCAAGAG 600
541 CAGAAGTCTGGTGGTGAATCTTTCGATGAACTCAAGAGAGGAGCTCCAGAGCAAGAG 600
601 GGAAGATTTGTTTATTAACCACTTACATCACTCAAGAGAGGAGGATGCAATCCGA 660
601 GGAAGATTTGTTTATTAACCACTTACATCACTCAAGAGAGGAGGATGCAATCCGA 660
661 CGCAGGGGGGGTGGAGAGCTGCAAGAGTGGGGGCTTTGGCATCTTCATTCATCCGTGG 720
661 CGCAGGGGGGGTGGAGAGCTGCAAGAGTGGGGGCTTTGGCATCTTCATTCATCCGTGG 720
721 CCTCTTCTCATCTAAGTCTCAACAAGTATTCAGGAATTAACAAGATGAGCGTGGCCA 780

721 CCTCTTCTCATCTAAGTCTCAACAAGTATTCAGGAATTAACAAGATGAGCGTGGCCA 780
781 AAATTCGAAGAGCTGTATTAACGTTGGAAGATGCAAGAAATGATGTCAAGATGGCTTCTC 840
781 AAATTCGAAGAGCTGTATTAACGTTGGAAGATGCAAGAAATGATGTCAAGATGGCTTCTC 840
841 ATGGAGTCAAAATTTGTCAATTCAGTAAAGATGGGGGAGCAAGACCTTACCCAGATCTGAT 900
841 ATGGAGTCAAAATTTGTCAATTCAGTAAAGATGGGGGAGCAAGACCTTACCCAGATCTGAT 900
901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGTACTGTCTCA 960
901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGTACTGTCTCA 960
961 GTGACATCTGACAGAGCTGGAGATTTGGCAGAGGTGCAATGATGATGAGGGGTGAGCCT 1020
961 GTGACATCTGACAGAGCTGGAGATTTGGCAGAGGTGCAATGATGATGAGGGGTGAGCCT 1020
1021 TTATATCATGGAGAGCACTCTCACTTAATAAGATCTTGGGCTGCTCCAAAGAGACTC 1080
1021 TTATATCATGGAGAGCACTCTCACTTAATAAGATCTTGGGCTGCTCCAAAGAGACTC 1080
1081 TGCAGCTGTGCTCTGAGACTGCAAGAGACAGAGTGAAGTTGTGCTTCCAGTATATTC 1140
1081 TGCAGCTGTGCTCTGAGACTGCAAGAGACAGAGTGAAGTTGTGCTTCCAGTATATTC 1140
1141 AGTTACCAAGATTAATTTTCCCACTACAGTCTGTGATGAGATCTGACCCAGAGAACT 1200
1141 AGTTACCAAGATTAATTTTCCCACTACAGTCTGTGATGAGATCTGACCCAGAGAACT 1200
1201 TCTTACCACATGGGCTGCAATTAATCTGAGAGTGAAGAGCCCAAGGCTCATATGAGAGAG 1260
1201 TCTTACCACATGGGCTGCAATTAATCTGAGAGTGAAGAGCCCAAGGCTCATATGAGAGAG 1260
1261 TTTAGAGCTGTGCTGAGAGCTTCAATTAATCACTAGTCTGAGCCATGAGAGAGAGAG 1320
1261 TTTAGAGCTGTGCTGAGAGCTTCAATTAATCACTAGTCTGAGCCATGAGAGAGAGAG 1320
1321 ACATCAACTTTTGAATCAAGAGTGAAGTGCCTGAGAGCACTGATCTTGAATGATTAATCA 1380
1321 ACATCAACTTTTGAATCAAGAGTGAAGTGCCTGAGAGCACTGATCTTGAATGATTAATCA 1380
1381 AGTATTTCTTCTTCCATCACTCCACAGGAGACACATGATCTGATGATGATCAAGAGCA 1440
1381 AGTATTTCTTCTTCCATCACTCCACAGGAGACACATGATCTGATGATGATCAAGAGCA 1440
1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
1501 TGCCTGCTAGGCTCTTAAGAAACAGTAAGAAAGAAAGCTTCAATGCTTCTGCGCAGGATC 1560
1501 TGCCTGCTAGGCTCTTAAGAAACAGTAAGAAAGAAAGCTTCAATGCTTCTGCGCAGGATC 1560
1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCAATCAATATTTGATCAATTTCACTTTCA 1620
1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCAATCAATATTTGATCAATTTCACTTTCA 1620
1621 AAGCAAACTCTATTTATGCTTTCTGTATTAATCTTTGATTAATCTTTCAAAATCTCT 1680
1621 AAGCAAACTCTATTTATGCTTTCTGTATTAATCTTTGATTAATCTTTCAAAATCTCT 1680
1681 GATTTCTGAAAAAAGAAATCAATTCCTCCCTCCCTCCACCAATGAAATCAACATATGTTA 1740
1681 GATTTCTGAAAAAAGAAATCAATTCCTCCCTCCCTCCACCAATGAAATCAACATATGTTA 1740
1741 GGGATTCAGTGGGGGATTTCTTATATCACTCTTAATAAATGTTGTTTCACTTTTAA 1800
1741 GGGATTCAGTGGGGGATTTCTTATATCACTCTTAATAAATGTTGTTTCACTTTTAA 1800
1801 AGTAAACACTTAATTAATTTTGGAGATCTCTGAAAAAATTTTAAAAA 1851

Db 1801 AGTAAACATTAAATTTTGGAGATCTGTGAAAAA 1851

RESULT 2
BC020689 1928 bp mRNA linear PRI 22-JAN-2002
DEFINITION Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418
IMAGE:4251802, mRNA, complete cds.
ACCESSION BC020689
VERSION BC020689.1 GI:18088383
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1928)
Strausberg, R.
Direct Submission
Submitted (03-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAL Plates: 35 Row: f Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7706386.
Location/Qualifiers
1. 1928
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:10404"
/db_xref="taxon:9606"
/clone="MGC:22418 IMAGE:4251802"
/issue_type="Skin, melanotic melanoma, high MDR."
/clone_id="NIH MGC_62"
/lab_host="DH10B"
/note="vector: pDNR-LIB"
167.1585
/codon_start=1
/product="plasma glutamate carboxypeptidase"
/protein_id="AAH20689.1"
/db_xref="GI:18088384"
/translation="MKRLIFAFKFGVHLISCSGKAIKRNKISRTFEIKERIASCG
DVAKALINLAVYGAQNRSTYERLALVDITGPRISGKNLEKAIQIOTNQDGLK
VHLKPRVPHMERGESAVMLEPRILKILIGLSSIGTPEGITAVALVVTSDLEK
RASEARGKIVVNPYINYSRTVOYTGQVNAKAGALASLRVASFSYSPHTG
IOEVQDVPKIPACITVEDAEMMSRASHGIKIVIQKMGAKTYPTDSFNTVAIT
GSKYRGVAVYVSHLSDMDVGQMDGGGAFISWEALSLIKDGLPKRLRLVMT
AERGGVGAFOYVQLKVNISNYLWESMDGTLPLPTGLQPTSEKRALMEPWSLI
QPLATITVLSHSGTDINFMIOAGVPAGASLIDLDLYKFFHSHSGDTTMDPKNNV
AAAVWAVSVYVADMEMLPRS"

BASE COUNT 554 a 435 c 466 g 473 t
ORIGIN

Query Match 99.4%; Score 1839.8; DB 9; Length 1928;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1844; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 GGATGAGCGGACCTTAGTCTCTGGAGACCGCTCTCGTCCGCGGTACAGAGCCGCCCTA 60

Db 69 GGCCTAGCCCGAGCTTAGTCTCTGGAGACCGCTCTCGTCCGCGGTACAGAGCCGCCCTA 128
QY 61 TCGATTATCTTAAACAAGAAACCAACCTGAAAAAATTAATTCCTATCTGGCAT 120
Db 129 TCGATTATCTTAAACAAGAAACCAACCTGAAAAAATTAATTCCTATCTGGCAT 188
QY 121 TTTTGGGTGTGTACCTTTTATCCCTGTGCTCTGGAAAGCTATATGCAAAATGGCA 180
Db 189 TTTTGGGTGTGTACCTTTTATCCCTGTGCTCTGGAAAGCTATATGCAAAATGGCA 248
QY 181 TCTTAAAGAGACTTTTGAAGAAATTAAGAAATATAGCCACCTGTGAGATGTGCTA 240
Db 249 TCTTAAAGAGACTTTTGAAGAAATTAAGAAATATAGCCACCTGTGAGATGTGCTA 308
QY 241 AAGCAATCAACCACTGAGTCTTTATATGCTTAAACCCAGAACATATCTATGACGATTTG 300
Db 309 AAGCAATCAACCACTGAGTCTTTATATGCTTAAACCCAGAACATATCTATGACGATTTG 368
QY 301 CACTTCTGTGTATCTGTGTGAGACCCAGACTGAGTGGCTCCAAGAACTTGAAGAAAGCCA 360
Db 369 CACTTCTGTGTATCTGTGTGAGACCCAGACTGAGTGGCTCCAAGAACTTGAAGAAAGCCA 428
QY 361 TCGAATTTATGTAACAAACCTGACAGACAGATGAGTGGAGAAAGTTTACCTGAGAGCCAG 420
Db 429 TCGAATTTATGTACAAACCTGACAGACAGATGAGTGGAGAAAGTTTACCTGAGAGCCAG 488
QY 421 TGAAGATACCCCACTGGGAGAGAGGAGAGAAATCAAGTGTGATGCTGAGACCAAGATTC 480
Db 489 TGAAGATACCCCACTGGGAGAGAGGAGAGAAATCAAGTGTGATGCTGAGACCAAGATTC 548
QY 481 ATTAAGATAGCCATCTGGAGTCTTGGAGAGCAATTTGGGACTCTCCAGAAAGGCAATTACAG 540
Db 549 ATTAAGATAGCCATCTGGAGTCTTGGAGAGCAATTTGGGACTCTCCAGAAAGGCAATTACAG 608
QY 541 CAGAAGTTCTGTGTGAGTACCTCTTTCGATGATGACAGAGAGGCTCAGAAAGCAAGAG 600
Db 609 CAGAAGTTCTGTGTGAGTACCTCTTTCGATGATGACAGAGAGGCTCAGAAAGCAAGAG 668
QY 601 GGAAGATTTGTTTATTAACCAACCTTACATCACTCAATCAAGAGCGTGCATACCGAA 660
Db 669 GGAAGATTTGTTTATTAACCAACCTTACATCACTCAATCAAGAGCGTGCATACCGAA 728
QY 661 CGCAGGGGGCGGTGGAAGCTGCCAAGGTGGGGGCTTTGGCATCTCATTCGATCCGTGG 720
Db 729 CGCAGGGGGCGGTGGAAGCTGCCAAGGTGGGGGCTTTGGCATCTCATTCGATCCGTGG 788
QY 721 CCTCTTCTCATCTACAGTCTCTACACAGATATTCAAGAAATCAAGATGCGTGGCCA 780
Db 789 CCTCTTCTCATCTACAGTCTCTACACAGATATTCAAGAAATCAAGATGCGTGGCCA 848
QY 781 AATTCCACAAGCCTGTATTACGCTGAGAGATGACAGAAATGATGTCAAGAAATGGCTTTC 840
Db 849 AGATTCCACAAGCCTGTATTACGCTGAGAGATGACAGAAATGATGTCAAGAAATGGCTTTC 908
QY 841 ATGGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGACCTACCAGATCTGATTT 900
Db 909 ATGGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGACCTACCAGATCTGATTT 968
QY 901 CTTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTTGTACTGTCA 960
Db 969 CTTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTTGTACTGTCA 1028
QY 961 GTGACACTGTGACAGCTGGAGATGTTGGCAGAGGTGTCATGATGATGCGGTGGAGCCT 1020
Db 1029 GTGACACTGTGACAGCTGGAGATGTTGGCAGAGGTGTCATGATGATGCGGTGGAGCCT 1088
QY 1021 TTATATGATGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCAAGAGAGCACTC 1080
Db 1089 TTATATGATGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCAAGAGAGCACTC 1148
QY 1081 TGGCGTGTGTCTCTGAGACTGAGAGAAACAAGGTGAGTTGTGCTTCAGATATTATC 1140

Db	1149	IGCGGCTGGTGCCTGGA	CTGAGAAAGAA	CAAGGTAAGTGGTGCCTTCAGATATTATC	1208
QY	1141	AGTTACACAAGTTAAATATTTTCCAACTACAGTCTGGTAGTAGAGTCTGACGCAAGACCT			1200
Db	1209	AGTTACACAAGTTAAATATTTTCCAACTACAGTCTGGTAGTAGAGTCTGACGCAAGACCT			1268
QY	1201	TCTTACCACTGGGGGTGCAATTCACTGGCAGTGAAGAAAGGCCAGGGCCATTCATGAGAGAG			1266
Db	1269	TCTTACCACTGGGGGTGCAATTCACTGGCAGTGAAGAAAGGCCAGGGCCATTCAGAGAGAG			1322
QY	1261	TTATAGCCTGCTGCAAGCCCTTCAATATCACTCAGATCCTGAGCATGAGAGAGAGCAG			1320
Db	1329	TTATAGCCTGCTGCAAGCCCTTCAATATCACTCAGATCCTGAGCATGAGAGAGAGCAG			1388
QY	1321	ACATCAACTTTTGGATCCAAAGCTGAGTGCCTGAGCCAGTCTACTGATGA	CTTATACA		1386
Db	1389	ACATCAACTTTTGGATCCAAAGCTGAGTGCCTGAGCCAGTCTACTGATGA	CTTATACA		1448
QY	1381	AGTATTTCTTCTTCATCATCTCCACGGAGACACATACCTGTCATGATCCAAAGCAGA			1444
Db	1449	AGTATTTCTTCTTCATCATCTCCACGGAGACACATACCTGTCATGATCCAAAGCAGA			1508
QY	1441	TGAATGTGTGCTGCTGTTTGGGGCTGTTGTTCTTATGTTGTTCAGACATGGAAGAAA			1500
Db	1509	TGAATGTGTGCTGCTGTTTGGGGCTGTTGTTCTTATGTTGTTCAGACATGGAAGAAA			1566
QY	1501	TGCTGCCTAGGCTCTAGAAA	CAGTAAGAAAGAAAGTTTCATGCTTCGGCAGGAATC		1566
Db	1569	TGCTGCCTAGGCTCTAGAAA	CAGTAAGAAAGAAAGTTTCATGCTTCGGCAGGAATC		1628
QY	1561	CTGGGCTGCAACTTTGGAAAACTCCTCTTCA	CATAACAATTCATCCATTCATCTTCA		1622
Db	1629	CTGGGCTGCAACTTTGGAAAACTCCTCTTCA	CATAACAATTCATCCATTCATCTTCA		1688
QY	1621	AAGCACACTCTAATTCATGCTTCTGTATATATCTTTCTGATATCTTCCAAATTCCT			1680
Db	1689	AAGCACACTCTAATTCATGCTTCTGTATATATCTTTCTGATATCTTCCAAATTCCT			1748
QY	1681	GATTCAGAAAAAGGAATCATTTCTCCCTCCCTCCACACATAGAAATCAACATATGTA			1740
Db	1749	GATTCAGAAAAAGGAATCATTTCTCCCTCCCTCCACACATAGAAATCAACATATGTA			1808
QY	1741	GGGATTAACAGTGGGGGCAATTTCTTTATATACACTCTTTAAAAACATGTTTCCATTTAA			1800
Db	1809	GGGATTAACAGTGGGGGCAATTTCTTTATATACACTCTTTAAAAACATGTTTCCATTTAA			1868
QY	1801	AGTAAACACTTAATTAATTTTTTGAAGATCTCTGAAAAAATTTTTTTTTT			1851
Db	1869	AGTAAACACTTAATTAATTTTTTGAAGATCAAAAAAATTTTTTTTTT			1919

LOCUS	BD127520	1660 bp	DNA	linear	PAT 18-SEP-2002
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.				
ACCESSION	BD127520				
VERSION	BD127520.1	GI:23222465			
KEYWORDS	JP 2002017375-A/2951.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1660)				
AUTHORS	Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.				
TITLE	Primer for synthesizing full-length cDNA and use thereof				
JOURNAL	Patent: JP 2002017375-A 2951 22-JAN-2002;				
COMMENT	HELIX RESEARCH INSTITUTE				
	OS	Homo sapiens (human)			
	PN	JP 2002017375-A/2951			
	PD	22-JAN-2002			
	PF	07-UTL-2000 JP 2000253172			

		PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO	
		PI ISHII,	
		PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI	
		SHINICHI KOJIMA,	
		PI TETSUJI OTSUKI, HISASHI KOGA	
		PC	
		C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/	
		10,	
		PC C12P1/02, C12Q1/68//C12P1/08, G06F17/30, C12N15/00, C12N5/00 CC	
		Primer for synthesizing full-length cDNA and use thereof FH Key	
		Location/Qualifiers	
		(129).. (1544).	
FEATURES		FT CDS	
source		Location/Qualifiers	
		1..1860	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
BASE COUNT		515 a 423 c 451 g 471 t	
ORIGIN			
Query Match		98.3%; Score 1820.4; DB 6; Length 1860;	
Best Local Similarity		99.7%; Pred. No. 0;	
Matches 1824; Conservative		0; Mismatches 6; Indels 0; Gaps 0	
QY	1	GGCTTAGCGCGGAGCTTAGTCTCTGGGAGCCGCTCCCTGCCGCCGTCAAGCCGCCCTTA	60
DB	31	GGCTTAGCGCGGAGCTTAGTCTCTGGGAGCCGCTCCCTGCCGCCGTCAAGCCGCCCTTA	90
QY	61	TCAGATTATCTTAACAAGAAAAACAACCTGSAAAAAAAAAATGAAATTCTTATCTTGCAAT	120
DB	91	TCAGATTATCTTAACAAGAAAAACAACCTGSAAAAAAAAAATGAAATTCTTATCTTGCAAT	150
QY	121	TTTTGGTGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGCA	180
DB	151	TTTTGGTGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGCA	210
QY	181	TCTCTAAGAGACTTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGAGATGTTGCTTA	240
DB	211	TCTCTAAGAGACTTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGAGATGTTGCTTA	270
QY	241	AAGCAATCATCAACCTAGCTGTTTATGTGAAGCCAGAACAGATCCTATGAGCGATTG	300
DB	271	AAGCAATCATCAACCTAGCTGTTTATGTGAAGCCAGAACAGATCCTATGAGCGATTG	330
QY	301	CACCTCTGTGTGATACCTGTGGAACCCAGACTGAGTGGCTCCAAAGAACTTAGAAAAAGCA	360
DB	331	CACCTCTGTGTGATACCTGTGGAACCCAGACTGAGTGGCTCCAAAGAACTTAGAAAAAGCA	390
QY	361	TCCAAATTATGTACCAAAAACCTGCAAGAGATGGCGCTGAGAAAGTTCACTCTGAGCCAG	420
DB	391	TCCAAATTATGTACCAAAAACCTGCAAGAGATGGCGCTGAGAAAGTTCACTCTGAGCCAG	450
QY	421	TGAGAAATACCCCACTGGGAGAGGGGAGAAAGATCAAGCTGTATGCTGAGCCCAAGAAATTC	480
DB	451	TGAGAAATACCCCACTGGGAGAGGGGAGAAAGATCAAGCTGTATGCTGAGCCCAAGAAATTC	510
QY	481	ATAAGATAGCCATCCTGGGCTCTTGCAAGACGATTTGGGACTCTCCAGAAAGCAATTACAG	540
DB	511	ATAAGATAGCCATCCTGGGCTCTTGCAAGACGATTTGGGACTCTCCAGAAAGCAATTACAG	570
QY	541	CAGAAATTCTGTGTGATCCTTTTGCATGAATGCAAGAAAGGGCTCTAGAAAGCAAGAG	600
DB	571	CAGAAATTCTGTGTGATCCTTTTGCATGAATGCAAGAAAGGGCTCTAGAAAGCAAGAG	630
QY	601	GGAAGATTGTTTATTAACAACCTTAACATCAATCACTCAAGAGACGGTGCAATACCGAA	660
DB	631	GGAAGATTGTTTATTAACAACCTTAACATCAATCACTCAAGAGACGGTGCAATACCGAA	690
QY	661	CGCAGGGGGCGGTGGAAGCTGCAAGAGTGGGGCTTTGGCATCTGTCAATTCGATCCGTGG	720
DB	691	CGCAGGGGGCGGTGGAAGCTGCAAGAGTGGGGCTTTGGCATCTGTCAATTCGATCCGTGG	750
QY	721	CCTCCTTCATCTACAGTCCCTCAACAAGATTAACAGAAATACAGATGGCGTCCCA	780

```

Db      751 CCTCTCTCCATCTACAGTCTCACAAGGATTCAGAAATCACAGAAATGCGCTGCCA 810
Qy      781 AAATTCACAAGCGCTGATTAACGGTGAAGATGCAAAATGCTGAAGATGGCTCTC 840
Db      811 AGATTCACAAGCGCTGATTAACGGTGAAGATGCAAAATGCTGAAGATGGCTCTC 870
Qy      841 ATGGGATCAAAATGCTGATTCAGTAAAGATGGGGGCAAGAACTTACCAGATCTGATT 900
Db      871 ATGGGATCAAAATGCTGATTCAGTAAAGATGGGGGCAAGAACTTACCAGATCTGATT 930
Qy      901 CTTTCAACACTGTAGCAGAGATCACTGGGACAAATATCCAGAAACAGTTGTACTGTCTCA 960
Db      931 CTTTCAACACTGTAGCAGAGATCACTGGGACAAATATCCAGAAACAGTTGTACTGTCTCA 990
Qy      961 GTGACACTGTGACAGCTGGGATTTGGGACAGGCTGCCATGAGATGCGGTGGACCT 1020
Db      991 GTGACACTGTGACAGCTGGGATTTGGGACAGGCTGCCATGAGATGCGGTGGACCT 1050
Qy      1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1080
Db      1051 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1110
Qy      1081 TGGGGCTGCTCTGTGACACTGACAGAAACAAGGTGAGTTGCTTCCAGTATTATC 1140
Db      1111 TGGGGCTGCTGCTGTGACACTGACAGAAACAAGGTGAGTTGCTTCCAGTATTATC 1170
Qy      1141 AGTTACCAAGGTAAATTTTCCACTACACTCTGTGTGATGAGATCTGACGAGAACT 1200
Db      1171 AGTTACCAAGGTAAATTTTCCACTACACTCTGTGTGATGAGATCTGACGAGAACT 1230
Qy      1201 TCTTACCACACTGGGCTGCAATTACTGGGACGTGAAAGGCGGACCATCATGAGAGAG 1260
Db      1231 TCTTACCACACTGGGCTGCAATTACTGGGACGTGAAAGGCGGACCATCATGAGAGAG 1290
Qy      1261 TTATGACCTCTGCTGACAGCCCTCAATATCACTCAAGTCTTGAGCCATGAGAGAGAG 1320
Db      1291 TTATGACCTCTGCTGACAGCCCTCAATATCACTCAAGTCTTGAGCCATGAGAGAGAG 1350
Qy      1321 ACATCAACTTTTGGATCCAACTGAGAGTGGCTGAGAGCCAGTCTTATGATACCTTATACA 1380
Db      1351 ACATCAACTTTTGGATCCAACTGAGAGTGGCTGAGAGCCAGTCTTATGATACCTTATACA 1410
Qy      1381 AGTATTTCTTCTTCATCACTCCACAGGAGACACATGACTGTCATGGATCCAAAGCAGA 1440
Db      1411 AGTATTTCTTCTTCATCACTCCACAGGAGACACATGACTGTCATGGATCCAAAGCAGA 1470
Qy      1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGGAAGAA 1500
Db      1471 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGGAAGAA 1530
Qy      1501 TGTGCTCCTAGGTCTAGAAACAGTAAAGAAAGAAAGCTTTTCACTGCTTGGCCAGGAATC 1560
Db      1531 TGTGCTCCTAGGTCTAGAAACAGTAAAGAAAGAAAGCTTTTCACTGCTTGGCCAGGAATC 1590
Qy      1561 CTGGGCTGCGAACTTTGGAAACCTCCTCTTCAATTAATTAATTAATTAATTAATTAAT 1620
Db      1591 CTGGGCTGCGAACTTTGGAAACCTCCTCTTCAATTAATTAATTAATTAATTAATTAAT 1650
Qy      1621 AAGCAACAATCTAATTTCACTGCTTCTGTTATTAATCTTCTGATTAATCTTCCAAATTCCT 1680
Db      1651 AAGCAACAATCTAATTTCACTGCTTCTGTTATTAATCTTCTGATTAATCTTCCAAATTCCT 1710
Qy      1681 GATTCTAGAAAAAGAAATCAATTCCTCCCTCCCTCCACACATGATCAACATATGAGTA 1740
Db      1711 GATTCTAGAAAAAGAAATCAATTCCTCCCTCCCTCCACACATGATCAACATATGAGTA 1770
Qy      1741 GGGATTACAGTGGGGGCAATTTCTTATTAATCACTTAAACCAATGTTTCCACTTTAA 1800
Db      1771 GGGATTACAGTGGGGGCAATTTCTTATTAATCACTTAAACCAATGTTTCCACTTTAA 1830
Qy      1801 AGTAAACACTAATTAATTTTGAAGATC 1830

```

```

Db      1831 AGTAAACACTAATTAATTTTGAAGATC 1860
RESULT 4
AK075132
LOCUS
DEFINITION
AK075132 1860 bp mRNA linear PRI 03-SEP-2002
Homo sapiens cDNA FLJ90651 fis, clone PLACE1004482, moderately
similar to Rattus norvegicus hematopoietic lineage switch 2 related
protein (H182-tp) mRNA.
ACCESSION
AK075132.1 GI:22761022
VERSION
Oligo capping, file (full insert sequence).
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Isogai,T., Ota,T., Nishikawa,T., Hayaashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Satou,K., Yamamoto,U., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masubo,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1860)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yama, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
source
1. 1860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004482"
/tissue_type="placenta"
/clone_lib="PLACE1"
/notes="Cloning vector: pME18SFL3"
129..1547
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC11423.1"
/db_xref="GI:22761023"
/translation="MKRLIFAFRGVHLISLCSGKAIKNGISKRTFEEIEEIEASCG
DVKKAILINLVYKRAQNSYERLALVDYVGPRLSGKNGKLEKAIQIYQWLQODGLEK
VHLEPVRIPIHMERGESAVMLEPRIRKIALILGSSGKTPPEIGITAEVLVTSFDEIQ
RRASPRARKIVVNOPINYSRTVOYRTOGVAEAAKVGALASLIRSVASISYSPHQ
IOXYODVPRKIPTRACTIVDPAEMSMASGIRIVIOIKGATVPEPTDSFNTVARTT
GSKYRPEOVVLSGHLSWDVGGANDGGATFISWELSLIKDIGLRPKRTLRVLMT
AEEQGGVAFQYVQLHKNVINSYSLVNESDAGFLPLTGLQFTSEKRAIAVEEVSIL
QPNLTIOVLSHGEGTIDINFWIOAGVPRASILLDDLYKFFPHSHSGDMTWDPKQNV
AAAVMAVSVYVADMEEMLPRS"
BASE COUNT
515 a 423 c 451 g 471 t
ORIGIN
Query Match 98.3%; Score 1820.4; DB 9; Length 1860;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1824; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy      1 GCGTAGGCGCGGAGCTTAAGTCCTGGGAGCGGCTCGCCGCGCTCAGAGCGCGCCTTA 60
Db      31 GCGTAGGCGCGGAGCTTAAGTCCTGGGAGCGGCTCGCCGCGCTCAGAGCGCGCCTTA 90
Qy      61 TCAGATTATCTTAACAAGAAACCAACTGGAAGAAAGAAATGAAATTCCTTATCTTCGAT 120

```

```

Db      91 TCAGATATCTTAACAGAAAAACAATGGAATAATCCCTTATCTTCGAT 150
Qy      121 TTTTGGGAGGTGTACCTTTATCCCTGTGCTGGGAAACCTTATGCAAGATGCA 180
Db      151 TTTTGGGAGGTGTACCTTTATCCCTGTGCTGGGAAACCTTATGCAAGATGCA 210
Qy      181 TCTTAAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTTGCTA 240
Db      211 TCTTAAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTTGCTA 270
Qy      241 AAGCAATCATCACTAGCTGTTTATGTTAAGCCAGAAACGATCTTATGAGCATTTG 300
Db      271 AAGCAATCATCACTAGCTGTTTATGTTAAGCCAGAAACGATCTTATGAGCATTTG 330
Qy      301 CACTCTGCTGATCTGTGGAACCCAGACGATGCTCCCAAGAACTTGAAGAAACCA 360
Db      331 CACTCTGCTGATCTGTGGAACCCAGACGATGCTCCCAAGAACTTGAAGAAACCA 390
Qy      361 TCCAAATTATGTAACCAAACTGCAAGCAAGATGGCTGGAGAAAGTTCACTGGAGCCAG 420
Db      391 TCCAAATTATGTAACCAAACTGCAAGCAAGATGGCTGGAGAAAGTTCACTGGAGCCAG 450
Qy      421 TGAGAAATCCCACTGGAGAGAGAGAGAAATCACTGTGTGATGCTGAGCCAGAAATTC 480
Db      451 TGAGAAATCCCACTGGAGAGAGAGAGAAATCACTGTGTGATGCTGAGCCAGAAATTC 510
Qy      481 ATAAATAGCCATCTCTGGGTCTTTGGAGAGATGGGACTCTTCAAGAAAGCATTAACG 540
Db      511 ATAAATAGCCATCTCTGGGTCTTTGGAGAGATGGGACTCTTCAAGAAAGCATTAACG 570
Qy      541 CAGAAGTCTGGTGTGACCTCTTGTGATGATGATGATGATGATGATGATGATGATG 600
Db      571 CAGAAGTCTGGTGTGACCTCTTGTGATGATGATGATGATGATGATGATGATGATG 630
Qy      601 GGAAGATTGTTTATTAACAACCTTACATCACTCAAGAGAGGCTGCAATACGAA 660
Db      631 GGAAGATTGTTTATTAACAACCTTACATCACTCAAGAGAGGCTGCAATACGAA 690
Qy      661 CGCAGAGGCGGTGGAAGCTGCAAGAGGCTTTGGCATCTTCAATCCGATCCGAG 720
Db      691 CGCAGAGGCGGTGGAAGCTGCAAGAGGCTTTGGCATCTTCAATCCGATCCGAG 750
Qy      721 CCTCCTTCTCATCTACAGTCTCTCAACAGATGATGATGATGATGATGATGATGATG 780
Db      751 CCTCCTTCTCATCTACAGTCTCTCAACAGATGATGATGATGATGATGATGATGATG 810
Qy      781 AAATTCACAGAGCTGTATTAAGGTGGAAGATGCAAGAAATGATGATGATGATGATG 840
Db      811 AGATTCACAGAGCTGTATTAAGGTGGAAGATGCAAGAAATGATGATGATGATGATG 870
Qy      841 ATGGATCAAAATTTGTCAATGCTTAAGATGAGGAGCAAAAGCTTAACCATTAATGAT 900
Db      871 ATGGATCAAAATTTGTCAATGCTTAAGATGAGGAGCAAAAGCTTAACCATTAATGAT 930
Qy      901 CCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGGTTGACTGTCA 960
Db      931 CCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGGTTGACTGTCA 990
Qy      961 GTGACATCTGACAGCTGGAGATGTTGGGAGAGGCTGCAATGATGATGATGATGATG 1020
Db      991 GTGACATCTGACAGCTGGAGATGTTGGGAGAGGCTGCAATGATGATGATGATGATG 1050
Qy      1021 TTATATCATGGAGACCTCTCACTTATTAAGATCTTGGGCTGCTGCAAAAGAGACTC 1080
Db      1051 TTATATCATGGAGACCTCTCACTTATTAAGATCTTGGGCTGCTGCAAAAGAGACTC 1110
Qy      1081 TGCGGCTGTGCTGTGAGCTGCAAGAAACAAAGGTGAGTGTGCTTCCAGATTAATTC 1140
Db      1111 TGCGGCTGTGCTGTGAGCTGCAAGAAACAAAGGTGAGTGTGCTTCCAGATTAATTC 1170
Qy      1141 AGTTACACAGAGTAATATTTTCAACTACAGCTGCTGTGATGAGATGATGATGATGATG 1200

```

```

Db      1171 AGTTACACAGAGTAATATTTTCCAACTACAGTCTGTGATGAGATGATGATGATGATG 1230
Qy      1201 TCTTACCACATGGGCTGAATTTCACTGGCAGGTGAAAGGCGCAGGCGCATCATGAGAGAG 1260
Db      1231 TCTTACCACATGGGCTGAATTTCACTGGCAGGTGAAAGGCGCAGGCGCATCATGAGAGAG 1290
Qy      1261 TTATGAGCTGTGCTGAGCCCTCAATATCACTCAGGTCTGAGCCATGAGAGAGAGAGAG 1320
Db      1291 TTATGAGCTGTGCTGAGCCCTCAATATCACTCAGGTCTGAGCCATGAGAGAGAGAGAG 1350
Qy      1321 ACATCAACTTTTGGATTCAGAGCTGAGTGTGAGAGCAGTCTTCTGATGATTAATCA 1380
Db      1351 ACATCAACTTTTGGATTCAGAGCTGAGTGTGAGAGCAGTCTTCTGATGATTAATCA 1410
Qy      1381 AGTATTTCTTCTTCATCACTCCAGAGAGACCATGATGATGATGATGATGATGATGATG 1440
Db      1411 AGTATTTCTTCTTCATCACTCCAGAGAGACCATGATGATGATGATGATGATGATGATG 1470
Qy      1441 TGAATGTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTGTTGACAGATGAGAGAA 1500
Db      1471 TGAATGTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTGTTGACAGATGAGAGAA 1530
Qy      1501 TGCTGCTAGGCTCTTAAAGACATGAAAGAAAGGTTTCTGATGCTTCTGAGCAATC 1560
Db      1531 TGCTGCTAGGCTCTTAAAGACATGAAAGAAAGGTTTCTGATGCTTCTGAGCAATC 1590
Qy      1561 CTGGGCTGCAACTTTGAGAAACCTCTTCACTAATCAATTCATCAATTCATCTTCA 1620
Db      1591 CTGGGCTGCAACTTTGAGAAACCTCTTCACTAATCAATTCATCAATTCATCTTCA 1650
Qy      1621 AAGCAACCTCTATTTGATGCTTCTGTTATTTATCTTCTGATGATGATGATGATGATG 1680
Db      1651 AAGCAACCTCTATTTGATGCTTCTGTTATTTATCTTCTGATGATGATGATGATGATG 1710
Qy      1681 GATTTGAAAAAGAAATCATTTCTCCCTCTCCACACATGATGATGATGATGATGATG 1740
Db      1711 GATTTGAAAAAGAAATCATTTCTCCCTCTCCACACATGATGATGATGATGATGATG 1770
Qy      1741 GGGATTCAGTGGGGGATTTCTTATATGACCTTAAAGAACTGTTTCACTTTAA 1800
Db      1771 GGGATTCAGTGGGGGATTTCTTATATCACTTTAAAGAACTGTTTCACTTTAA 1830
Qy      1801 AGTAACACTTAATTAATTTTGGAGATC 1830
Db      1831 AGTAACACTTAATTAATTTTGGAGATC 1860

RESULT 5
AX006440
LOCUS      AX006440      1860 bp      DNA      linear      PAT 06-SEP-2000
DEFINITION Sequence 5 from Patent WO0004157.
ACCESSION  AX006440
VERSION     AX006440.1  GI:994575
KEYWORDS
SOURCE
ORGANISM   Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
            Neeffs,J.M., Peeters,D.C. and Pangalos,M.
            Cloning and characterisation of novel mammalian peptidases
            Patent: WO 0004157-A 5 27-JAN-2000;
            JANSSEN PHARMACEUTICA NV (BE); NEEFS JEAN MARC EDMOND FERNAND (BE);
            PEETERS DANIELLE CELINE GEORGE (BE); PANGALOS MENILAS (GB)
FEATURES
            source
            1. 1860
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT  518 a      420 c      451 g      471 t
ORIGIN
Query Match      96.3%; Score 1783.2; DB 6; Length 1860;

```


carboxypeptidase/prostate-specific membrane antigen
 JOURNAL J. Biol. Chem. 274 (17), 11742-11750 (1999)
 MEDLINE 9923495
 PUBMED 10206990
 2 (bases 1 to 1794)
 REFERENCES Gnjirras, R., Richard, C., El-Alfy, M., Morales, C. R., Potier, M. and
 AUTHORS Pehnezetsky, A. V.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUN-1999) Medical Genetics, Sainte-Justine Hospital,
 Montreal University, 3175 Cote Sainte-Catherine, Montreal, QU H3T
 1C5, Canada

FEATURES

source

Location/Qualifiers

1..1794
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8q22.2; between D8S257 and D8S270"
 1..1794
 /gene="PGCP"
 25..1650
 /gene="PGCP"
 /EC_number="3.4.17.21"
 /function="endopeptidase"
 /note="member of M28 peptidase family; similar to
 glutamate carboxypeptidase II and to
 N-acetyl-aspartyl-alpha-glutamate carboxypeptidase
 (NALADase); similar to sequences with GenBank Accession
 Numbers R18560 and G07404; prostate-specific membrane
 antigen (PSMA)"
 /codon_start=1
 /product="blood plasma glutamate carboxypeptidase
 precursor"
 /protein_id="AAD31418.1"
 /db_xref="GI:4877698"

gene

CDS

sig_peptide
 /gene="PGCP"
 157..1647
 /gene="PGCP"
 /product="blood plasma glutamate carboxypeptidase"
 205..207
 /gene="PGCP"
 /note="putative; glycosylation site"
 559..561
 /gene="PGCP"
 /note="putative; glycosylation site"
 1081..1083
 /gene="PGCP"
 /note="putative; glycosylation site"
 misc_feature
 /note="putative; glycosylation site"
 1090..1092
 /gene="PGCP"
 /note="putative; glycosylation site"
 1210..1212
 /gene="PGCP"
 /note="putative; glycosylation site"
 BASE COUNT 522 a 422 g 457 t 2 others
 ORIGIN

Query Match 95.0%; Score 1758.6; DB 9; Length 1794;
 Beac Local Similarity 99.7%; Pred. No. 0;
 Matches 1772; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

75 CAAGAAACCAACTGGAAAGAAATGAATCTTACTTCGATTTTTCGGTGTGT 134.

|||||
 Db 1 CAAGAAACCAACTGGAAAGAAATGAATCTTACTTCGATTTTTCGGTGTGT 60
 135 CACCTTTATTCCTGTCCTCTGGGAAAGCTATATGACAAATGGCATCTCTAAGAGACT 194
 Db 61 CACCTTTATTCCTGTCCTCTGGGAAAGCTATATGACAAATGGCATCTCTAAGAGACT 120
 195 TTGAAGAAATTAAGAAAGAAATAGCAGCTGTGGAGATGTGTGTAAGCAATCATCAAC 254
 Db 121 TTGAAGAAATTAAGAAAGAAATAGCAGCTGTGGAGATGTGTGTAAGCAATCATCAAC 180
 255 CTACCTGTTATATGTTAAAGCCCAAGACAGATCTATAGAGGATTTGGCATTTGTGTTAT 314
 Db 181 CTACCTGTTATATGTTAAAGCCCAAGACAGATCTATAGAGGATTTGGCATTTGTGTTAT 240
 315 ACTGTTGACCCAGACAGTGAAGTGTCCAAAGAACTTAAGAAAGCAATCCAAATATATAC 374
 Db 241 ACTGTTGACCCAGACAGTGAAGTGTCCAAAGAACTTAAGAAAGCAATCCAAATATATAC 300
 375 CAAGAACTGACGAAAGATGGGCTGGAGAAAGTTCACTGGAGCCAGTGAAGATACCCAC 434
 Db 301 CAAGAACTGACGAAAGATGGGCTGGAGAAAGTTCACTGGAGCCAGTGAAGATACCCAC 360
 435 TGGAGAGGGGAGAAAGATCAGCTGTATGCTGGAGCCAAAGATTTCAATAGATACCATC 494
 Db 361 TGGAGAGGGGAGAAAGATCAGCTGTATGCTGGAGCCAAAGATTTCAATAGATACCATC 420
 495 CTGGGCTCTTGACAGACATTTGGGATCTCTCCAAAGGCATTAACGAGAAAGTTCTGTG 554
 Db 421 CTGGGCTCTTGACAGACATTTGGGATCTCTCCAAAGGCATTAACGAGAAAGTTCTGTG 480
 555 GTGACCTCTTTCGATGTAAGTGCAGAGAGGGCCCTCAGAGCAAGAGGAAAGATGTTGTT 614
 Db 481 GTGACCTCTTTCGATGTAAGTGCAGAGAGGGCCCTCAGAGCAAGAGGAAAGATGTTGTT 540
 615 TATAACCAACCTTACATCACTCACTCAAGACGCTGCAATACCAACGAGGGGCGGTG 674
 Db 541 TATAACCAACCTTACATCACTCACTCAAGACGCTGCAATACCAACGAGGGGCGGTG 600
 675 GAAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCATTTGATGCTGTCCTCTTCCATC 734
 Db 601 GAAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCATTTGATGCTGTCCTCTTCCATC 660
 735 TACAGTCTCACAAGATTAATCAGAAATACAGAGATGGCGCCCAAAATTCACAGAGCC 794
 Db 661 TACAGTCTCACAAGATTAATCAGAAATACAGAGATGGCGCCCAAGATTCACAGAGCC 720
 795 TGTATTAAGTGAAGATGCAAGAAATGATCAAGATGGCTTCTCATGGGATCAAAAT 854
 Db 721 TGTATTAAGTGAAGATGCAAGAAATGATCAAGATGGCTTCTCATGGGATCAAAAT 780
 855 GTCATTCAGCTAAAGATGGGGGCAAGAACTTACCAATCTGATTTCTTCAACACTGTA 914
 Db 781 GTCATTCAGCTAAAGATGGGGGCAAGAACTTACCAATCTGATTTCTTCAACACTGTA 840
 915 GCAGAGATCACTGGAGCAAAATATTCAGAAACAGGTTGATCGGACAGTGGACATCTGAC 974
 Db 841 GCAGAGATCACTGGAGCAAAATATTCAGAAACAGGTTGATCGGACAGTGGACATCTGAC 900
 975 AGCTGGATGTTGGGACAGGTCAGATGATGATGATGATGATGATGATGATGATGATGAT 1034
 Db 901 AGCTGGATGTTGGGACAGGTCAGATGATGATGATGATGATGATGATGATGATGATGAT 960
 1035 GCATCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTTGGCGGCTGTGCTC 1094
 Db 961 GCATCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTTGGCGGCTGTGCTC 1020
 1095 TGACCTGCAAGAAAGAAAGTGGAGTGTGGCTTCCAGATTAATCAAGATTAACAGGTA 1154
 Db 1021 TGACCTGCAAGAAAGAAAGTGGAGTGTGGCTTCCAGATTAATCAAGATTAACAGGTA 1080
 1155 AATATTTTCAACTCACTGTGTGATGAGATGATGAGATGATGAGATGATGAGATGATGAG 1214

Dp	1081	AAATTTTCAACTACAGTCTGTATGGAAGTCTGACGAGAAACCTTTACCACTGCG	1141									
Qy	1215	CTGCATTTACTGCGGACGTGAAAGGCCAGGGCCATCATGAGAGAGTTATGAGCTGCTG	1274									
Dp	1141	CTGCATTTTCACTGGCAGTGAAGAAAGGCCAGGGCCATCATGAGAGAGTTATGAGCTGCTG	1200									
Qy	1275	CAGCCCTCAATATCATCTCAGGTCTGTGCGCATGGAGAAAGGACAGATCATGCTTTGG	1334									
Dp	1201	CAGCCCTCAATATCATCTCAGGTCTGTGCGCATGGAGAAAGGACAGATCATGCTTTGG	1266									
Qy	1335	ATCCAGCTGGAGTGCCTGAGCCAGTCTACTTGAATGATCTTAATCAAGTATTTCTTCTC	1394									
Dp	1261	ATCCAGCTGGAGTGCCTGAGCCAGTCTACTTGAATGATCTTAATCAAGTATTTCTTCTC	1320									
Qy	1395	CATCATCTCCACGGAGACACATGACTGTATCATGATCCAAAGACAGATGTTGCTGCT	1454									
Dp	1321	CATCATCTCCACGGAGACACATGACTGTATCATGATCC - CAGCAGATGATGTTGCTGCT	1379									
Qy	1455	GCTGTTTGGGGCTGTTGTTCTTAATGTTGTGACAGCATGGAAGAAAGTCGCTAAGTCC	1514									
Dp	1380	GCTGTTTGGGGCTGTTGTTCTTAATGTTGTGACAGCATGGAAGAAAGTCGCTAAGTCC	1439									
Qy	1515	TAGAAACAGTAAAGAAAGAAAGTTTTCATGCTTCTGCGCAGAGAAATCTGGGCTGCACT	1574									
Dp	1440	TAGAAACAGTAAAGAAAGAAAGTTTTCATGCTTCTGCGCAGAGAAATCTGGGCTGCACT	1499									
Qy	1575	TTGGAACCTCCTCTTCAATACATTAATTCATTCATCTTCAAGCAACTCTAT	1634									
Dp	1500	TTGGAACCTCCTCTTCAATACATTAATTCATTCATCTTCAAGCAACTCTAT	1559									
Qy	1635	TTCATGCTTCTGTATATATCTTTCTGTATCTTCCAAATTTCTGTATTTCTAGAAAAG	1694									
Dp	1560	TTCATGCTTCTGTATATATCTTTCTGTATCTTCCAAATTTCTGTATTTCTAGAAAAG	1619									
Qy	1695	GAATCATTTCTCCCTCCCTCCACACATAGATCAATAPATGGAGGATTAACAGTGG	1754									
Dp	1620	GAATCATTTCTCCCTCCCTCCACACATAGATCAATAPATGGAGGATTAACAGTGG	1679									
Qy	1755	GGCATTTCTTATATACCTCTTAAAAACATTTGTTCCACTTAAAGTAACCTTAT	1814									
Dp	1680	GGCATTTCTTATATACCTCTTAAAAACATTTGTTCCACTTAAAGTAACCTTAT	1739									
Qy	1815	AAATTTTGGAAAGATCTCTGAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1851									
Dp	1740	AAATTTTGGAAAGATCTCTGAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1776									
RESULT 7	BD105816	Secretory protein.	1778 bp	DNA	linear	PAT 18-SEP-2002						
LOCUS	BD105816	Secretory protein.										
DEFINITION	BD105816	Secretory protein.										
ACCESSION	BD105816	Secretory protein.										
VERSION	BD105816.1	GI:23200634										
KEYWORDS	JP 2002502234-A/7.											
SOURCE	JP 2002502234-A/7.											
ORGANISM		synthetic construct										
		synthetic construct										
		artificial sequences.										
		1 (bases 1 to 1778)										
REFERENCE		Jacob, K., McCoy, J.M., Racie, L.A., Lavallie, E.R., Merberg, D. and Spaulding, V.										
AUTHORS		Jacob, K., McCoy, J.M., Racie, L.A., Lavallie, E.R., Merberg, D. and Spaulding, V.										
TITLE		Secretory protein										
JOURNAL		Patent: JP 2002502234-A 7 22-JAN-2002;										
COMMENT		GENETICS INSTITUTE INC. PN JP 2002502234-A/7 PD 22-JAN-2002 PF 16-APR-1997 JP 1997537384 PR 18-APR-1996 US 08/634325, 13-JAN-1997 US 08/783520 PI KENNETH JACOBS, JOHN M MCCOY, LISA A RACIE, EDWARD R LAVALLIE, PI DAVID MERBERG, PI WIKKI SPAULDING PC C12N15/12, C07K14/47, A61K38/17 CC Strandedness: Double; CC Topology: Linear;										

FEATURES	FH	Key	Location/Qualifiers	Location/Qualifiers
source			1..1778	/organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630"
BASE COUNT		514 a	386 c	422 g 456 t
ORIGIN				
Query Match		94.6%;	Score 1751;	DB 6; Length 1778;
Best Local Similarity		99.4%;	Pred. No. 0;	
Matches 1757;		Conservative	0;	Mismatches 10; Indels 0; Gaps 0
OY	85	AACTGGAAAAAATGAAATTCCTTATCTTGGCATTTTGGGTGTGTCACTTTTAT	144	
DB	3	AGCTTGGCAGAGATGAAATTCCTTATCTTGGCATTTTGGGTGTGTCACTTTTAT	62	
OY	145	CCCTGTGCTCTGGGAAAGCTATATGCAAGAAATGATCTCTTAAGAGACCTTTGAAGAA	204	
DB	63	CCCTGTGCTCTGGGAAAGCTATATGCAAGAAATGATCTCTTAAGAGACCTTTGAAGAA	122	
OY	205	TAAAGAGAAATATCCAGCTGTGAGATGTTGCTTAAAGCAATCATCAACTGCTGTT	264	
DB	123	TAAAGAGAAATATGCGACCTGTGAGATGTTGCTTAAAGCAATCATCAACTGCTGTT	182	
OY	265	ATGTGAAGCCCAAGACGATCTTATGAGCATTTGGCATTTCTGTTGATATCTGTTGAC	324	
DB	183	ATGTGAAGCCCAAGACGATCTTATGAGCATTTGGCATTTCTGTTGATATCTGTTGAC	242	
OY	325	CCAGACTAGTGGCTCCCAAGAACTTAAAGAAAGCCATCAATATATGACCAAAACCTGC	384	
DB	243	CCAGACTAGTGGCTCCCAAGAACTTAAAGAAAGCCATCAATATATGACCAAAACCTGC	302	
OY	385	AGCAAGATGGGCTGAGAAAGTTCACTCTGGACCCAGTGAATATCCCACTGGAGAGGG	444	
DB	303	AGCAAGATGGGCTGAGAAAGTTCACTCTGGACCCAGTGAATATCCCACTGGAGAGGG	362	
OY	445	GAGAAAGATACGCTGTGATGCTGGAGCCAAAGATTCATTAAGATGACCTGGGTCTTG	504	
DB	363	GAGAAAGATACGCTGTGATGCTGGAGCCAAAGATTCATTAAGATGACCTGGGTCTTG	422	
OY	505	GCAGCAGCATTTGGGACCTCTCCAGAAAGGATATACAGCAGAAATTCGTGTGACCTCTT	564	
DB	423	GCAGCAGCATTTGGGACCTCTCCAGAAAGGATATACAGCAGAAATTCGTGTGACCTCTT	482	
OY	565	TCGATGAATCGACAGAAAGGCTCTGAAAGCAAGAGGGAAATGTTGTTTATTAACCAAC	624	
DB	483	TCGATGAATCGACAGAAAGGCTCTGAAAGCAAGAGGGAAATGTTGTTTATTAACCAAC	542	
OY	625	CTTACATTAACCTACCAAGAGCGTGCAATACGAAACGAGGGGCGGTGGAAAGCTGCCA	684	
DB	543	CTTACATTAACCTACCAAGAGCGTGCAATACGAAACGAGGGGCGGTGGAAAGCTGCCA	602	
OY	685	AGGTGGGGGCTTTGGCATCTCTCATTTGCAATCCGAGGCTCTTCTTCATCTACAGTCTTC	744	
DB	603	AGGTGGGGGCTTTGGCATCTCTCATTTGCAATCCGAGGCTCTTCTTCATCTACAGTCTTC	662	
OY	745	ACACAGGATTTTCAGGAATATCCAGATGGCGTCCCAAAATTCOAACAGCTGTATTAACG	804	
DB	663	ACACAGGATTTTCAGGAATATCCAGATGGCGTCCCAAAATTCOAACAGCTGTATTAACG	722	
OY	805	TGGAAGATGCAAGAAATGATGTCAAAAGGGCTTCTCATGGGATCAAAATTTGTCATTACG	864	
DB	723	TGGAAGATGCAAGAAATGATGTCAAAAGGGCTTCTCATGGGATCAAAATTTGTCATTACG	782	
OY	865	TAAAGATGGGGCAAAAGACTTACCAGATCTGATTTCTTCAACACTGTAGCAGAGATCA	924	
DB	783	TAAAGATGGGGCAAAAGACTTACCAGATCTGATTTCTTCAACACTGTAGCAGAGATCA	842	
OY	925	CTGGAGAGAAATATTCAGAAACAGGTTGTATCTGTCAATGAGCAATCTGACACGTGGGAATG	984	
DB	843	CTGGAGAGAAATATTCAGAAACAGGTTGTATCTGTCAATGAGCAATCTGACACGTGGGAATG	902	

```

Oy 985 TTGGGACAGGCTGCGCATGATGATGCGGCGGACCTTTATATCATGGAGACACTCTCAC 1044
Db 903 TTGGGACAGGCTGCGCATGATGATGCGGCGGACCTTTATATCATGGAGACACTCTCAC 962
Oy 1045 TTATTTAAAGATCTTGGGCTGCGCTCCAAAGAGACTCTGCGGCTGGTGTCTGGACTGCGAG 1104
Db 963 TTATTTAAAGATCTTGGGCTGCGCTCCAAAGAGACTCTGCGGCTGGTGTCTGGACTGCGAG 1022
Oy 1105 AAGAAACAAGGTGGAGTGGTGGCTTCCAGTATATATCATGTTACACAAAGTAAATATTTTCCA 1164
Db 1023 GAGAAACAAGGTGGAGTGGTGGCTTCCAGTATATATCATGTTACACAAAGTAAATATTTTCCA 1082
Oy 1165 ACTACAGTCTGGTGTGGAGTCTGACGACGAGAACCTTTCTTACCAGCTGGGCTGCACATTTCA 1224
Db 1083 ACTACAGTCTGGTGTGGAGTCTGACGACGAGAACCTTTCTTACCAGCTGGGCTGCACATTTCA 1142
Oy 1225 CTGGCAGTGAAGAAAGGCCAGGGCCATCATGAGAGAGTTATGAGCCTGTGACGCCCTCA 1284
Db 1143 CTGGCAGTGAAGAAAGGCCAGGGCCATCATGAGAGAGTTATGAGCCTGTGACGCCCTCA 1202
Oy 1285 ATATCACTCAGGTCTGAGCCATGAGAGAGGACACATCACTTTTGGATCCAGCTG 1344
Db 1203 ATATCACTCAGGTCTGAGCCATGAGAGAGGACACATCACTTTTGGATCCAGCTG 1262
Oy 1345 GAGTCCCTGAGACCACTCTACTTGTATGACTTATACAGATATTTCTTCTCCATCACTCCC 1404
Db 1263 GAGTCCCTGAGACCACTCTACTTGTATGACTTATACAGATATTTCTTCTCCATCACTCCC 1322
Oy 1405 ACGGAGACACCACTGCTGATGATCCAAAGCAAGTGAATGTTGCTGCTGCTGTTGGG 1464
Db 1323 ACGGAGACACCACTGCTGATGATCCAAAGCAAGTGAATGTTGCTGCTGCTGTTGGG 1382
Oy 1465 CTGTTGTTCTTATGTTGTGAGACATGAGAGAAATGCTGCTGCTGCTGCTGAGAAACGT 1524
Db 1383 CTGTTGTTCTTATGTTGTGAGACATGAGAGAAATGCTGCTGCTGCTGCTGAGAAACGT 1442
Oy 1525 AAGAAAGAAGCTTTTCACTGCTTGGGACAGAAATCTGCGGTCTGCAACTTTGAGAAACT 1584
Db 1443 AAGAAAGAAGCTTTTCACTGCTTGGGACAGAAATCTGCGGTCTGCAACTTTGAGAAACT 1502
Oy 1585 CCTCTTCAATTAACATTTTCACTCAATTCATCTTCAAGACACACTTATTTTCACTCTT 1644
Db 1503 CCTCTTCAATTAACATTTTCACTCAATTCATCTTCAAGACACACTTATTTTCACTCTT 1562
Oy 1645 CTGTTATATATCTTTCTTGAATCTTCTTCAAAATCTCTGATCTTGAAGAAAGAAATCTCT 1704
Db 1563 CTGTTATATATCTTTCTTGAATCTTCTTCAAAATCTCTGATCTTGAAGAAAGAAATCTCT 1622
Oy 1705 CCCCTCCCTCCACACATGATGATCAATGATGAGGATTAACAGTGGGSCATTTCTT 1764
Db 1623 CCCCTCCCTCCACACATGATGATCAATGATGAGGATTAACAGTGGGSCATTTCTT 1682
Oy 1765 TATATCACTCTTTAAACAAATTTTCACTTTAAAGTAAACCTTAAATATTTTGG 1824
Db 1683 TATATCACTCTTTAAACAAATTTTCACTTTAAAGTAAACCTTAAATATTTTGG 1742
Oy 1825 AAGATCTCTGAAGAAAAAATTTTAAAAA 1851
Db 1743 AAGATCTCTGAAGAAAAAATTTTAAAAA 1769

```

```

RESULT 8
AF107834 1472 bp mRNA linear PRI 12-UTR-1999
LOCUS AF107834
DEFINITION Homo sapiens clone LCHI aminopeptidase mRNA, complete cds.
ACCESSION AF107834
VERSION AF107834.1 GI:5442029
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1472)

```

```

AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.
TITLE Cloning of the human aminopeptidase gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1472)
AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
Sinica, RM 433, 128, Yen-Chun-Yuan Road SEC 2, Taipei 11529, Taiwan
FEATURES
Source
1..1472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LCHI"
/tissue_type="liver"
/dev_stage="fetus"
54..1472
/feature="secretory protein"
/codon_start=1
/product="aminopeptidase"
/protein_id="P46321.1"
/db_xref="GI:5442030"
/translation="MKFLIFAFPGGHLISCSKAIKNGISKRTREBIKEIASCG
/DNAKAIINLAVYGAQNSYERLALVDIVGPRLSGKNLEKATQIYQWQDDLEK
VHLEPVI PHMERGESBAMLEPRIRHIALIGLSSISGTEPEGITAEVLVTSFDEQ
RRASEARGKIVVYNQPIYNSRTQGAVEAKGALASLIRSVAFSISYSPHG
IOEYDQVPIPRACITVEDABMSBMSHGIKIVIQKGAATYPTDSFNTVAEIT
GSKRPEQVVLVSGHSDMSVQCGAMDGGAFLISWEALSLIKDLGKPRTLRLVMT
AEKGGVGAFOYQIHKVNIISNYSLVMSBNGFTLPGLQPTSGEKRAIMEBMSLL
QPLNITGVLSHGEGTDINFMIQAGVPAASLDDLYKYFFPHSHGDTMTWMDPKNNV
AAAVWAVSVVVDMEEMLPSS"
BASE COUNT 411 a 325 c 381 g 355 t
ORIGIN
Query Match 79.2%; Score 1465.6; DB 9; Length 1472;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1468; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 46 GTGAGAGCGCGCCATGATGATTTATTAACAGAAACCAACTGGAGAAATGAAT 105
Db 1 GTGAGAGCGCGCCATGATGATTTATTAACAGAAACCAACTGGAGAAATGAAT 60
Oy 106 TCCTTATCTTGCAATTTTCCGTTGATGATCACTTTTATCCCTGTGCTCTGGAGAACTA 165
Db 61 TCCTTATCTTGCAATTTTCCGTTGATGATCACTTTTATCCCTGTGCTCTGGAGAACTA 120
Oy 166 TATGCAAGATGCAATCTTAAGAGACTTTTGAAGAAATTAAGAAATTAAGCACT 225
Db 121 TATGCAAGATGCAATCTTAAGAGACTTTTGAAGAAATTAAGAAATTAAGCACT 180
Oy 226 GTGAGAGTGTGTAAGCAATCATCAACCTAGCTGTTATGTAAGCCAGAAAGAT 285
Db 181 GTGAGAGTGTGTAAGCAATCATCAACCTAGCTGTTATGTAAGCCAGAAAGAT 240
Oy 286 CCTATGAGCGATTTGGCACTTCTGTTGATGATCTTGAAGCCAGACTGATGATG 345
Db 241 CCTATGAGCGATTTGGCACTTCTGTTGATGATCTTGAAGCCAGACTGATGATG 300
Oy 346 ACTTGAAGAAAGCCATCAATTTATGTAAGAAACCTGACAGAAATGAGTGGAGAAAG 405
Db 301 ACTTGAAGAAAGCCATCAATTTATGTAAGAAACCTGACAGAAATGAGTGGAGAAAG 360
Oy 406 TTCACTGGAGCGATGAGATATCCCACTGGGAGAGGAGAGAAATCAGCTGATGC 465
Db 361 TTCACTGGAGCGATGAGATATCCCACTGGGAGAGGAGAGAAATCAGCTGATGC 420
Oy 466 TGAAGCAAGAAATCATTAAGATAGCAATCTGGGCTTGGAGAGAGCAATTTGGACTCTC 525
Db 421 TGAAGCAAGAAATCATTAAGATAGCAATCTGGGCTTGGAGAGAGCAATTTGGACTCTC 480
Oy 526 CAGAAAGCATTAACAGCAAGATTTCTGTTGATGATCTTTTGAATGAACTGACAGAAAGG 585
Db 481 CAGAAAGCATTAACAGCAAGATTTCTGTTGATGATCTTTTGAATGAACTGACAGAAAGG 540

```

QY	586	CCTCAGAAAGCAAGGGGAAAGTGTGTTTATTAACAACCTTACATCAACTACGAA	645
Dp	541	CCTCAGAAAGCAAGGGGAAAGTGTGTTTATTAACAACCTTACATCAACTACGAA	600
QY	646	CGGTGCATTAACGAACGAGGGGGGCGGTGGAAGCTGCACAGAGGTGGGGCTTGTGCATCTC	705
Dp	601	CGGTGCATTAACGAACGAGGGGGGCGGTGGAAGCTGCACAGAGGTGGGGCTTGTGCATCTC	660
QY	706	TCATTCCATCCGTGGCTTCCTTCTTCATCTTACAGTCTTCAACAGATATTCAGAAATAC	765
Dp	661	TCATTCCATCCGTGGCTTCCTTCTTCATCTTACAGTCTTCAACAGATATTCAGAAATAC	720
QY	766	AGATGGGGTGGCCCAAAATTCACAACGCTGTATTAACGTGGGAABAATCAGAAATGATGT	825
Dp	721	AGATGGGGTGGCCCAAAATTCACAACGCTGTATTAACGTGGGAABAATCAGAAATGATGT	780
QY	826	CAGAATGGCTTCTTCATGGAGATCAAAATTTGATTCACAGTAAAGATGGGGGCAAGACCT	885
Dp	781	CAGAATGGCTTCTTCATGGAGATCAAAATTTGATTCACAGTAAAGATGGGGGCAAGACCT	840
QY	886	ACCAGATATCTGATTCCTTCAACACTGTACAGAGATCACTGGAGCAAAATATCCAGAAC	945
Dp	841	ACCAGATATCTGATTCCTTCAACACTGTGTACAGAGATCACTGGAGCAAAATATCCAGAAC	900
QY	946	AGGTGTGACTGTGTCAGTGGACATCTGGACAGCTGGGAGTGGGGAGGGTGCATGATG	1005
Dp	901	AGGTGTGACTGTGTCAGTGGACATCTGGACAGCTGGGAGTGGGGAGGGTGCATGATG	960
QY	1006	ATGGCGGTGAGCCTTTATATCATGGGAAGCACTCACTTATTAAGATCTTGGGCTGC	1065
Dp	961	ATGGCGGTGAGCCTTTATATCATGGGAAGCACTCACTTATTAAGATCTTGGGCTGC	1020
QY	1066	GTCGAAGAAGGACTCTGCGGCTGGTGTCTGTGATCTGCAGAGAACAAGATGGAGTTGGTG	1125
Dp	1021	GTCGAAGAAGGACTCTGCGGCTGGTGTCTGTGATCTGCAGAGAACAAGATGGAGTTGGTG	1080
QY	1126	CCTTCCAGATATTAACAGTTACAAAGGTAATATTTCCAATCAACAGCTGTGGATGATGT	1185
Dp	1081	CCTTCCAGATATTAACAGTTACAAAGGTAATATTTCCAATCAACAGCTGTGGATGATGT	1140
QY	1186	CTGACGAGGAAACCTTCTTATCCCACTGGGCTGCATTAATCACTGGCGAGTAAAGGCCAGGG	1245
Dp	1141	CTGACGAGGAAACCTTCTTATCCCACTGGGCTGCATTAATCACTGGCGAGTAAAGGCCAGGG	1200
QY	1246	CCATCATGGAAGAGTTATGAGCTGTGCAAGCCCTCAATATCACTCAGGTCTTGAGCC	1305
Dp	1201	CCATCATGGAAGAGTTATGAGCTGTGCAAGCCCTCAATATCACTCAGGTCTTGAGCC	1260
QY	1306	ATGGAGAAGGACAGACATCAACTTTTGGATTCGAAGCTGAGAGTGCCTGGAAGCAATCAAC	1365
Dp	1261	ATGGAGAAGGACAGACATCAACTTTTGGATTCGAAGCTGAGAGTGCCTGGAAGCAATCAAC	1320
QY	1366	TTGATGACTTATTAACAAGTATTTCTTCCATCACTCCACGAGAGACAACATGACTGTCA	1425
Dp	1321	TTGATGACTTATTAACAAGTATTTCTTCCATCACTCCACGAGAGACAACATGACTGTCA	1380
QY	1426	TGATTCCAAGACAGATGATGTGTGCTGTGCTGTTTGGGCTGTGTTTCTTATGTTGTG	1485
Dp	1381	TGATTCCAAGACAGATGATGTGTGCTGTGCTGTTTGGGCTGTGTTTCTTATGTTGTG	1440
QY	1486	CAGACATGGAAGAAATGCTGCTGAGTGCCTGAG	1517
Dp	1441	CAGACATGGAAGAAATGCTGCTGAGTGCCTGAG	1472

RESULT 9	1796 bp	mRNA	linear	ROD 16-APR-2003
BC037067				
LOCUS				
DEFINITION	Mus musculus plasma glutamate carboxypeptidase, mRNA (cDNA clone			
	MGC:46857 IMAGE:498536), complete cds.			
ACCESSION	BC037067			
VERSION	BC037067.1	GI:22477497		

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

MGC.
Mus musculus (house mouse)
Mus musculus
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus
1 (Baes 1 to 1796)
1 Struhsberg, R. L., Feinsgold, E. A., Grouse, L. H., Derge, J. G.,

TITTLE

JOURN
MEDI T

REFERENCES

TITL

11

COMMENT

FEATURE

sources

gene

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
Series: IRAX Plate: 81 Row: g Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9055233.
Location/Qualifiers
1..1796
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:46857 IMAGE:4989536"
/tissue_type="Colon, normal. 5 month old male mouse."
/clone_id="NCI CGAP_Co24"
/lab_host="DH10B"
/note="vector: pCMV-SF0RT6"
1..1796

QY	131	TGTCACCCCTTTATCCCGTGTGGGAAAGCTATATGCAAGATGGCATCTCMAAG	190
Db	89	TGTCACCTTTTAAAGCTTGAGCTGTAAAGCTATTCAAGAATGGTGTTCACAG	148
QY	191	GACTTTGAAGAAATTAAGAAATATGCCAGCTGTGGAATGTGTGTAAGCAATAT	250
Db	149	AACATTTCGAAGAAATTAAGAAATATGCCCACTATTAAGATGTGTGTAAGCAATAT	208
QY	251	CAACCTAGCTTTATGGTAAAGCCCAACAGATCCTATGAGCGATTGGCACTTCTG	310
Db	209	CAACCTTGCTTTATGGTAAATACCAAGACGGTTCATATAGCGTTTGGGACTTCTAG	268
QY	311	TGATATCTTTGACCCAGACTGATGGCTCCAGAACCTAGAAAAGCATCCAAATAT	370
Db	269	TGATATCTTTGACCCAGACTGATGGCTCTAAGAACCTAGAAAAGCATTCAAATAT	328
QY	371	GTACCAAAACCTGCGCAAGATGGGCTGGAGAAATTCACCTGGAGCCAGTGAAGATAC	430
Db	325	GTACCAAAACCTGCGCAAGATGGGCTTGAAAATTTATCCTGGAGCGGTCAAGATACC	388
QY	431	CCACTGGAGAGGGAGAAAGATCAGCTGTGTATGCTGAGCCCAAGATTCATTAAGATAG	490
Db	385	CCACTGGAGAGGGAGAAAGATCTGCAGTATGCTTGAAGCTCGAATTCACAGATGGC	448
QY	491	CATCTGGGCTTTGGCAGCAGCATTTGGGATCTCTCCAGAAAGCATTAAGCAAGATCT	550
Db	445	TATCTAGGCTTTGGCAGCAGCATTTGGGATCTCCCAAGAGGCATCAAGCAAGATGCT	508
QY	551	GGTGTGACCTCTTTCGATGGAAGCTGCAAGAAAGGCTCAGAAAGCAAGGGAAGATGT	610
Db	505	GGTGTGACCTCTTTCGATGGAAGCTTCAAGAAAGGATCAAGAGCAAGGGAAGATAT	568
QY	611	TGTTTATTAACCAACTTACATCACTACTCAAGAGCGGTGCAATACCGAACGCAAGGGGC	670
Db	569	TGTTTATTAACCAAGCTTACACTGGCTTATGAGAAAGCTGTGCAAGTACCGGGGTGCAAGGGAGC	628
QY	671	GGTGAAGCTGCCAAGGTGGGGGCTTTGGCATCTCATTTGGATCCGTGGGCTCCTCTC	730
Db	625	TGTGAAGCTGCCAAGGTGGGAGCTGTGGCATCCTCTCATCAATATAGATGTTCTTTTC	688
QY	731	CATCTACAGTCTCAACAGGATTTCAAGAAATACAGAGATGGCGTCCCAAAATTCAC	790
Db	689	CATCTACAGTCTCTCAACAGGATTTCAAGAAATATCAAGATGGTGTGCCCAAGATTCAC	748
QY	791	AGCCTGTATTAAGGTGGAAGATGCAAGAAATGATGTCAAGAAATGGCTTCTCATGGGATCA	850
Db	749	AGCCTGTATTAAGGTGGAAGATGCAAGAAATATGATGTCAAGAAATGGCTTCTGTGGGAACA	808

QY	851	AATTGTCATTCAGCTAAAGATGGGGGAGAAAGACCTTACCAGATATCTGATTCCTTCAACAC	910
Db	809	AATTGTCATTCATCTGGAAATGGGAGCAAGAACCTTATCCAGATATCTGATTCCTTCAATAC	868
QY	911	TGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTGTACTGGTCAGTGGACATCT	970
Db	869	TGTTGCAGAGAGATCACTGGAGCAATGTATCCAGAGAAAGTTGTCTCTGGTCAAGTGAATTTT	928
QY	971	GGACAGCTGGGATSTTGGGGCAGGGTGGCAATGGATGATGGCGGTGGAGCCTTTATATCATG	1030
Db	929	GGACAGCTGGGATSTTGGGGCAGGGTGGCACTGGATGATGGTGGTGGAGCCTTCAATATCATG	988
QY	1031	GGAGCACYCTCACTTATTAAGATCTTGGGCTGGCTCCAAAGAGACTTGGCGGCTGT	1090
Db	989	GGAGCACYCTCACTTGTAAAGATCTTGGGCTGGCTCCAAAGAGAGACTCTGGCGGCTGTGT	1048
QY	1091	GCTTGGACCTGAGAAAGAACAAAGGTGGAGTTGGTGTCTTCCAGTATATATCAGTTACACAA	1150
Db	1049	GCTCTGGACCTGAGAAAGAACAAAGGAGAAATTTGGTGTCTTCCAGTATATATGAGCTACATPA	1108
QY	1151	GGTAAATATTTCCAACTACAGTCTGGTGTATGGAGTCTGACGCGAGGAACCTTCTTACCAC	1210
Db	1109	GGCAATATTTCCAAAGTACAGTCTGGTGTATGGAGTCTGACGCGAGGAACCTTCTTACCAC	1168
QY	1211	TGGGCTGCAATTCACCTGGCAGGTGAAGGAGCCAGGGCCATCATGGAGAGGTTATGACCT	1270
Db	1169	TGGACTGCAATTCACCTGGCAGGTGAAGGAGCCAGGGCCATCATGAGAGAGTCAATGAATCT	1228
QY	1271	GCTGCAGCCCCCTCAATATATCACTGAGGCTTGAGCCATGAGAGAGGAGCAGACATCACTT	1330
Db	1229	TCTGCAGCCCCCTCAATATATCACTGAGGCTTTTATGATGGAGAGAGAACTGTATATTAACCT	1288
QY	1331	TTGGATCCAAAGCTGGAGTGCCTGGAGCCAGTCTATGATGACTTATACAGATATTTCTT	1390
Db	1289	CTGGATCCAAAGCTGGAGTGCCTGGAGCCAGTCTGTGATGACTTATACAGATATTTCTT	1348
QY	1391	CTTGCATCATCTCCACGAGGAGCACATGATCTGTATATGGATTCCAAAGAGATGATGTGTC	1450
Db	1349	TTTCCATCATCTCCACGAGGAGCACATGATCTGTATATGGATTCCAAAGAGATGATGTGTC	1408
QY	1451	TGCTGCTGTGGGCTGTGTTTCTTATGTGTGTGACAGATGAGAGAAATGCTGCTAG	1510
Db	1409	TGCTGCTGTGGGCTGTGTTGCTTATGTGTGTGACAGATGAGAGAAATGCTGCCAG	1468
QY	1511	GTCTTAAAGACGTAAGAAAGAAACGTTTTCATGCTTCTGGCAGAGAAATCCTGGGTCTGC	1570
Db	1469	GTCTTAAAGAGAACAAAGAGAGAGACCTTGTCTCTGTGTGGAGATCCCAACTGTGA	1528
QY	1571	AACCTTGGAGAAACCTCCTTTCACATATCAATTTATCATCAATTCATCTTCAAAGACAACT	1630
Db	1529	ATCTTTCACACATCTCATCTGTCCAAAGAGGTATATCATTTATATCCACAGGGCATAGTT	1588
QY	1631	CTATTTCAATGCTTCTGTATATATCTTTTATGATCTTCCAAATTTCTGTGATTTAGAA	1690
Db	1589	TTCTTTATACCTTCTGTATATCACTTCTCTTATATACCTTC---TATCTGTTTCTAGAA	1645
QY	1691	AAAGATCAATTCCTCCCTCCCTCC	1716
Db	1646	TAAATCAATGATCCCTACTGCAACAC	1671

RESULT	10			
LOCUS	AX400829			
SEQUENCE	505	1778 bp.	DNA	linear
FROM	Patent WO0210453.			
AX400829				
AX400829.1	GI:21337009			
KEYWORDS				
SOURCE				
ORGANISM				
	<i>Rattus norvegicus</i> (Norway rat)			
	<i>Rattus norvegicus</i>			
	Euarchonta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			

REFERENCE 1
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and
TITLE Elashoff, M.R.
JOURNAL Molecular toxicology modeling.
Gene Patent: WO 0210453-A 505 07-FEB-2002;
FEATURES
source location/Qualifiers
1. 1778
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/note="EMBL/GenBank Accession No. AF097723"
BASE COUNT 488 a 391 c 442 g 457 t
ORIGIN
Query Match 61.9%; Score 1145; DB 6; Length 1778;
Best Local Similarity 81.9%; Pred. No. 2.4e-288;
Matches 1334; Conservative 0; Mismatches 290; Indels 5; Gaps 1;
Rattus.
73 AACAGAAACCACTGGAAGAAAAAATGAAATTCCTATCTTGGCATTTTGGGTGTG 132
95 AGCAAGAAAGAAAGAACTAGGACATAGAGTTCCTTCTCTGTTGCTGTG 154
133 TTCACTTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGATGSCATCTTAAGAGA 192
155 TTCACTTTTCTCTGGGCTCTGGAAGAACTATATACAGAGTGTGTTCTCAGCAA 214
193 CTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTGCTAAGCAATCATCA 252
215 CATTTCAAGAAATTAAGAAATAGCCAACTAAGAGATGTGCTAAGCAATTAATCA 274
253 ACCAGAGCTTTATAGTAAAGCCAGAACATCTATAGAGAGTTGGCATCTGCTGTG 312
275 ACCTGCTGTTATGGAATAACAGAACCGTCTATAGAGGTTTGGACTTCTAGTGTG 334
313 ATACTGTGAGACCACTAGTGTGCTCAAGAACTTGAAGAAAGCCATCCAAATATATG 372
335 ATACTGTGAGACCACTAGTGTGCTCAAGAACTTGAAGAAAGCCATCCAAATATATG 394
373 ACCAAACCTGACAGATGGGCTGGAAGAAAGTTACCTGAGGCACTGAGAAATACCC 432
395 ACCAAACCTGACAGATGGGCTGGAAGAAAGTTACCTGAGGCACTGAGAAATACCC 454
433 ACTGGAGAGGGGAGAAATACAGTGTGATGCTGTGAGACCAAGATTTCAATAGAGCA 492
455 ACTGGAGAGGGGAGAAATCTGCAATGATGTGTCTCTGAATTCACAAAGTTGGCTA 514
493 TCCTGGGCTTGGAGAGCAATGGGACTCTCCAGAAAGCACTTACAGAGAGTTCTGG 552
515 TTTTAGGCTTGGGAGAGCAATGGGACTCTCTCCAGAAAGCACTTACAGAGAGTTCTGG 574
553 TGTGACCTCTTGTGATGATGAGAGAGGGCTCAGAAAGCAAGAGAGAGATTTGTG 612
575 TGTGACCTCTTGTGATGATGAGAGAGGGCTCAGAAAGCAAGAGAGATTTGTG 634
613 TTTTATACCAACCTTACATCACTACTCAAGAGAGGCTGATACCAAGAGAGAGAGAG 672
635 TTTTATACCAACCTTACATCACTACTCAAGAGAGGCTGATACCAAGAGAGAGAGAG 694
673 TGAAGAGTGCAGAGAGGAGGCTTGGGCTCTCTATTCGATCCGTGGCTCTCTCTCA 732
695 TGAAGAGTGCAGAGAGGAGGCTTGGGCTCTCTATTCGATCCGTGGCTCTCTCTCA 754
733 TCTACAGTCTCTACAGAGATTTCAAGATACAGAGATGCGTCCAAATTTCCACAG 792
755 TCTACAGTCTCTACAGAGATTTCAAGATACAGAGATGCGTCCAAATTTCCACAG 814
793 CCTGATTAAGGTGAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGAT 852
815 CCTGATTAAGGTGAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGAT 874
853 TTGATATTACAGTAAAGATGGGGGCAAGACCTACCAAGATCTAGATTCCTTCAACAG 912

875 TTGATATTACAGTAAAGATGGAGCAAGACCTTACAGATACAGATTCCTTCAACAG 934
913 TACAGAGATACCTGGAGAGCAAAATTCAGAAAGGTTGATCTGATGAGATCTGG 972
935 TTGAGAGATACCTGGAGAGCAAAATTCAGAAAGGTTGATCTGATGAGATCTGG 994
973 ACAGCTGGAGATGTTGGGAGGAGTCCAGATGATGATGATGATGATGATGATGATGAT 1032
995 ACAGCTGGAGATGTTGGGAGGAGTCCAGATGATGATGATGATGATGATGATGATGATGAT 1054
1033 AAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGAGCTCTGGGCTGTGC 1092
1055 AAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGAGCTCTGGGCTGTGC 1114
1093 TCTGAGCTGCAAGAAACAGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 1152
1115 TCTGAGCTGCAAGAAACAGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 1174
1153 TAAATATTTCCAACTACAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1212
1175 CAATATTTCCAACTACAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1234
1213 GGTGCAATTCATCTGCAAGTGAAGAAAGCCAGAGCCATCATGAGAGAGTTAAGACCTGC 1272
1235 GGTGCAATTCATCTGCAAGTGAAGAAAGCCAGAGCCATCATGAGAGAGTTAAGACCTGC 1294
1273 TGCAGCCCTCTAATATCATCTAGTCTGAGCCATGAGAGAGAGAGAGAGAGAGAGAG 1332
1295 TGCAGCCCTCTAATATCATCTAGTCTGAGCCATGAGAGAGAGAGAGAGAGAGAGAG 1354
1333 GGATCCAGCTGAGAGTCTGAGAGCCAGTCTACTGATGATGATGATGATGATGATGATGAT 1392
1355 GGATCCAGCTGAGAGTCTGAGAGCCAGTCTACTGATGATGATGATGATGATGATGATGAT 1414
1393 TCCATCACTCCCAAGAGAGACCATGATCTGATGATGATGATGATGATGATGATGATGAT 1452
1415 TCCATCACTCCCAAGAGAGACCATGATCTGATGATGATGATGATGATGATGATGATGAT 1474
1453 CTGCTGTTGGGCTGTTGTTCTTATGTTGTTGAGAGATGAGAGAGAGAGAGAGAGAG 1512
1475 CTGCTGTTGGGCTGTTGTTCTTATGTTGTTGAGAGATGAGAGAGAGAGAGAGAGAG 1534
1513 CCTGAAACAGTAAAGAAAGAGATTTCAATGATCTTGGGCAAGATCTGGGCTGTGCA 1572
1535 CCTGAAACAGTAAAGAAAGAGATTTCAATGATCTTGGGCAAGATCTGGGCTGTGCA 1594
1573 TTTGGAAGAACTCTCTTCAATACATTAATTCATCAATTCATCAATTCATCAATTCAT 1632
1595 TTTGAG 1649
1633 ATTTGATGCTTCTGTTATATATCTTTTCTGATCTTTTCAATTCATCTGATCTTGA 1692
1650 GTTTCTTATATCTTCTGTTTAAACATCTTCTGATACGCTTTTAACTGTTTAA 1709
1693 AGGAATCAT 1701
1710 AGTATCAT 1718
RESULT 11
AF097723
LOCUS 1778 bp mRNA linear ROD 10-NOV-1998
DEFINITION Rattus norvegicus hematopoietic lineage switch 2 related protein
ACCESSION AF097723
VERSION AF097723
KEYWORDS AF097723.1 GI:3851631
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
TITLE
JOURNAL
1 (pages 1 to 1778)
Chen, Y. and Talmage, D.
Direct Submission
Submitted (07-OCT-1998) Institute of Human Nutrition, Columbia
University, 701 West 168th Street Room 5-503, New York, NY 10032,
USA


```

QY      1693 AGGATCAT 1701
Db      1710 AGTATCAT 1718

RESULT 12
AF131077      1726 bp mRNA linear ROD 03-JAN-2003
LOCUS      Rattus norvegicus liver annexin-like protein (LAL) mRNA, complete
DEFINITION
ACCESSION  AF131077
VERSION     AF131077.1 GI:7108712
KEYWORDS    Rattus norvegicus (Norway rat)
SOURCE      Rattus norvegicus
ORGANISM    Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 1726)
AUTHORS    Della Fazio,M.A., Piobico,D., Bartoli,D., Castellini,M.,
            Brancorsini,S., Viola Magni,M. and Servillo,G.
TITLE      lal-1, a differentially expressed novel gene during proliferation
            in liver regeneration and in hepatoma cells
JOURNAL    Genes Cells 7 (11), 1183-1190 (2002)
MEDLINE    22278398
PUBMED     12390252
REFERENCE   2 (bases 1 to 1726)
AUTHORS    Servillo,G., Della Fazio,M.A., Piobico,D., Bartoli,D.,
            Castellini,M., Brancorsini,S. and Viola Magni,M.
TITLE      Direct Submission
JOURNAL    Submitted (25-FEB-1999) Institute of General Pathology, University
            of Perugia, Policlinico Monteluce, Perugia 06100, Italy
FEATURES
source     1..1726
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            1..1726
            /gene="LAL"
            /gene="Lal"
            /note="Involved in liver regeneration; contains unique
            annexin domain at the N-terminus"
            /product="liver annexin-like protein"
            /protein_id="AAF36518.1"
            /db_xref="GI:7108713"
            /translation="MRLPFLFVAVVHLPSIGSKAIYKSGVSRFOEIKETANYE
            DVAKATINLAVYGYONRSYERGLVDTVGPRLSGSKNEKATQIOMYLOODGLEN
            VHLFOVRIPIHMERGESAVWVPRVHILALILGSGSICTPREGITATVAVLVASFEIQ
            RRASEAKGIVVYNQPTDYGYQYREKAVAEAKGAVASLIRVASSTISPHFG
            HOGVODVPKIPYACITTEDENMSRTASRDKIVHLKMGAKYPTDSEFNVAEIT
            GSKYPERVIVVSGHLDSDWDGQALDDGSAFSLWELSLVKDGLRPKTLRLVLT
            ABERGQVAGQYELHRAKANSKYSLVMEADSGTFLPTGLPTGSDKRAIMKEVMSLL
            OPLNTKVFNDAGTDIDINFWIOAGVPSALRDLVYKFFRHHSHDPTMTAMDPRQNVV
            AAVMVAIVAVVADMEEMLRPS"
BASE COUNT 474 a 385 c 428 g 439 t
ORIGIN
Query Match 61.8%; Score 1143.8; DB 10; Length 1726;
Best Local Similarity 80.9%; Pred. No. 4.9e-288;
Matches 1346; Conservative 0; Mismatches 312; Indels 5; Gaps 1;
QY      39 GCGCCGCTGACAGCCGCCCTATCAGATTATCTTAACAAGAAAACAAGTGAAGAAAAA 98
Db      3 CGGAGCTGCGGCTGCTCTGACAAACCTGTCACCAAGAAAGAAAGAACTAGGACA 62
QY      99 ATGAATTCCTTATCTTCGATTTTGGGTGGTTCACCTTTATCCCTGTCTGGG 158
Db      63 ATGAGGTTCTTCTTCCTGTTGTTGCTGTGTTCACCTTTTCTCTGGGCTCTGGA 122
QY      159 AAGCTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAAATTAAGAGATA 218

```

```

Db      123 AAAGCTATATACAAAGAGTGTGTTTCTCAGGAACATTTCAAGAAATTAAGAGATA 182
QY      219 GCGAGCTGTGAGATGTTTGTCTAAAGCAATCATCAACTAGTGTTTATGTAAAGCCAG 278
Db      183 GCCAACTATGAAGATGTTGTCTAAAGCAATTAACCTCTGTGTTATGAAAAATACAG 242
QY      279 AACGATCTATGAGCGATTGGCACTTCTGTTGTAATCTGTGGAGCCAGCTAGTGGC 338
Db      243 AACCGGTGCTATGAGCGTTTGGGACTTCTAGTTGATCTGTTGAGCCAGACTAGTGGC 302
QY      339 TCCAAAGACTTGAAGAAAAGCCATCCAAATTATGTACCAAACTGCAGCAAGATGGCTG 398
Db      303 TCTAAGAACTTAGAAGAAAGCTATCCAAATCATGTATGACAAAACCTGCAGCAAGATGGCTG 362
QY      399 GAGAAATTCACCTGAGGCCAGTGAAGTAATCCCATCTGGAGAGAGGAGAAAGATCACT 458
Db      363 GAAAACCTCCACCTGGAGCGAGGTCAAGTAATCTTCACCTGAGAGAGGCGCAAGATCTGCA 422
QY      459 GTGATGCTGAGGCCAAGATTCATTAAGTATGACCATCTGGGCTTTGGCAGAGCATTTGG 518
Db      423 GTGATGCTGAGGCCCTGCAATTCACAGTTGCTATTTAGCCCTTGGCGGAGCATTTGGG 482
QY      519 ACTCCTCCAGAAAGCATTACAGCAGAAAGTTCTGGTGTGACCTTTTCATGATGATGCTGAG 578
Db      483 ACTCCTCCTGAAGGTATCACAGCAGAAAGTACTGTGTGCTCTTTTGTGAACTTCAA 542
QY      579 AGAAGGCTCTAAGAGCAAGAGGAGAGATGTTGTTTATTAACCAACTTACATCACTAC 638
Db      543 AGAAGGCTCTAAGAGCAAGAGGAGAGATGTTGTTTATTAACCAACTTACATCACTAC 602
QY      639 TCAAGAGCGGTCAATCCGAGAGCGAGGCGGTGAGCTGCAAGTGGGGGCTTGG 698
Db      603 GCGAAACTGTGACGATCCGAGAGCGGAGCGGTGAGCTGCAAGTGGGGGCTTGG 662
QY      699 GCATCTCTCATTCGATCCGTGGCTCTCTCTCATCTACATGCTCTCACAGGATTTGAG 758
Db      663 GCATCTCTCATTCGATCAGTATGAGTCTCTTTTTCATCTACAGTCTCTCACAGGATTTGAG 722
QY      759 GAATACCAAGATGCGGTGCCCAAAATTCACACAGCTGTATTACGTTGAGAGATGCAGA 818
Db      723 GGAATACCAAGATGCGGTGCCCAAAATTCACACAGCTGTATTACGTTGAGAGATGCAGA 782
QY      819 ATGATGTCAGATATGCTTCTCATGAGTCAAAATGTCATTCAGCTTAAGATGGGGCA 878
Db      783 ATGATGTCAGATATGCTTCTCATGAGTCAAAATGTCATTCAGCTTAAGATGGGGCA 842
QY      879 AAGACTACCCAGATACCTGATTCCTTCAACACTGTACAGAGATCACTGGAGGCAATTTAT 938
Db      843 AAGACTATCCAGATACAGATTCCTTCAACACTGTACAGAGATCACTGGAGGCAATTTAT 902
QY      939 CCAGAACAGTTGATCTGTGACGTGACATCTGACAGCTGGAGATGTTGGGCGAGGTGCC 998
Db      903 CCAGAACAGTTGATCTGTGACGTGACATCTGACAGCTGGAGATGTTGGGCGAGGTGCC 962
QY      999 ATGATATAGCGGTGAGAGCTTTATATCATGAGAACACTCTCACTTATTAAGATCTT 1058
Db      963 CTGATATAGCGGTGAGAGCTTTATATCATGAGAACACTCTCACTTATTAAGATCTT 1022
QY      1059 GGGCTGCTCAAAAGAGACTGTGCGGCTGTCTCTGACTGACAGAAAGCAAGAGTGA 1118
Db      1023 GGGCTGCTCAAAAGAGACTGTGCGGCTGTCTCTGACTGACAGAAAGCAAGAGTGA 1082
QY      1119 GTTGTGCTCTCCAGATTAATTAAGTTACAGATTAATTAATTTCAACTACAGTGGTG 1178
Db      1083 GTTGTGCTCTCCAGATTAATTAAGTTACAGATTAATTAATTTCAACTACAGTGGTG 1142
QY      1179 ATGAGTCTGACCGAAGAACTTTCTTAACCACTGGGTGCAATTCATCTGCACTGGAAG 1238
Db      1143 ATGAGGCTGACCGAAGAACTTTCTTAACCACTGGGTGCAATTCATCTGCACTGGAAG 1202
QY      1239 GCGAGGCTCATGAGAGAGTTATGAGCTCTGAGCCCTCAATATCACTCAGGTC 1298

```

Db 1203 GCCAGGCTATCATGAGAAAGTCTGAGTCTCTGCAACCCCTCATATATGACCAAGGTC 1262
 Oy 1299 CTGAGCCATGAGAGGAGGACAGATCACTTTTGATCCAAAGTGGAGTCTGGAGCC 1358
 Db 1263 TTTATATGATGAGAGGAGGACATGACATTAATCTTGATCTCAAGCTGGAGTCTGGAGCC 1322
 Oy 1359 AGCTACTTGAATGACTTATACAGATATTTCTTCTTCATCACTCCACGAGACACCATG 1418
 Db 1323 AGTGGCAGATGATCTTACAGATATTTCTTTTCCATCACTCCATGAGAGACCATG 1382
 Oy 1419 ACTGTCATGATCCAAACAGATGATGTTGCTGCTGTTGGGCTGTTGTTCTTAT 1478
 Db 1383 ACTGCATGATCCAAACAGATGATGTTGCTGCTGTTGGGCTGTTGTTCTTAT 1442
 Oy 1475 GTTGTGAGACATGAGAAAGATGCTGCTAGTCTGTAAGAAAGTAAAGAAAGCTT 1538
 Db 1443 GTTGTGAGACATGAGAAAGATGCTGCTAGTCTGTAAGAAAGTAAAGAAAGCTT 1502
 Oy 1539 TTCAATGCTTGGCCAGAGATCTGGGCTGCAACTTTGGAAGAACTCCCTCTTCAATAC 1598
 Db 1503 TTGATCTGCAAGCTGGAGATCCCATTCGGGATTTTCAAGCAGCCATCTTCAAGAC 1562
 Oy 1595 AATTTCATCAATTCATTTCAAGCACTCTATTTTCATGCTTTCTGTTATATCTT 1658
 Db 1563 CTGTATATACATCAATCCCGTGGCACA-----GTTCTTTATACCTTCTGTTAACCAT 1617
 Oy 1655 CTGTATATCAATTCATCTGATTTCTGAGAAAGAAAGATCAT 1701
 Db 1616 CTTCCTGATACGCTTTTACCTGTTCTAGAAATGATATCAT 1660

RESULT 13
 AF107835 1596 bp mRNA linear ROD 12-JUL-1999
 LOCUS
 DEFINITION Mus musculus clone LCH-M1 aminopeptidase mRNA, complete cds.
 ACCESSION AF107835
 VERSION AF107835.1 GI:5442031
 KEYWORDS

SOURCE
 ORGANISM
 Mus musculus (house mouse)

REFERENCE
 AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1596)
 Liu, C.H., Lin, B.Y. and Chang, L.Y.
 TITLE Cloning of the mouse aminopeptidase gene
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 1596)
 Liu, C.H., Lin, B.Y. and Chang, L.Y.
 TITLE Direct Submission
 JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
 Sinica, RM 433,128, Yen-Chiun-Yuan Road SEC 2, Taipei 11529, Taiwan

FEATURES
 source
 1..1596
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="LCH-M1"
 /tissue_type="liver"
 184..1596
 /note="secretory protein"
 /codon_start=1
 /product="aminopeptidase"
 /protein_id="A0D43215.1"
 /db_xref="GI:5442032"
 /translation="MRSFLFLFIYHLIALSGKAVPKNGVSQRTFREIKEELANYEDV
 AKALINLAIVGKYQNSRYERGLLVTVGPRLSGSKLLEKAIQMYONLQODGLENVH
 LEQVRIPIHMERGESAVTLEBRILKMAILIGSSILGPPGSLTEVLVAVSFDELQIR
 ASEAKRIIVNQGTEYKTYOYRVOGAEEAAGVAVASLIGVAFSISPSPTGQR
 KYODGPKIIPACTIVDEAMSMASRKIVTHLMEMAKYPTDPSFNTVAEITS
 MYPEVTVSGHLDMDVGGALDDGGAGATISWALSLVNDLGRPRRTIRLVIMTBS
 EGGSTGASQYIEHLKANKISKSLVMEADSGFLPTGLQFTGSDKARALIMKVMNLQD
 LNTKVPNSNEGTDINFIQAGVPGASLRDLVLYKFFPHSHSGDTVMQPKQNVAA
 AVAAVAVVAVDMDMLPRS"

CDS

/note="secretory protein"
 /codon_start=1
 /product="aminopeptidase"
 /protein_id="A0D43215.1"
 /db_xref="GI:5442032"
 /translation="MRSFLFLFIYHLIALSGKAVPKNGVSQRTFREIKEELANYEDV
 AKALINLAIVGKYQNSRYERGLLVTVGPRLSGSKLLEKAIQMYONLQODGLENVH
 LEQVRIPIHMERGESAVTLEBRILKMAILIGSSILGPPGSLTEVLVAVSFDELQIR
 ASEAKRIIVNQGTEYKTYOYRVOGAEEAAGVAVASLIGVAFSISPSPTGQR
 KYODGPKIIPACTIVDEAMSMASRKIVTHLMEMAKYPTDPSFNTVAEITS
 MYPEVTVSGHLDMDVGGALDDGGAGATISWALSLVNDLGRPRRTIRLVIMTBS
 EGGSTGASQYIEHLKANKISKSLVMEADSGFLPTGLQFTGSDKARALIMKVMNLQD
 LNTKVPNSNEGTDINFIQAGVPGASLRDLVLYKFFPHSHSGDTVMQPKQNVAA
 AVAAVAVVAVDMDMLPRS"

BASE COUNT 415 a 357 c 421 g 403 t
 ORIGIN
 Query Match 60.4%; Score 1118.4; DB 10; Length 1596;
 Best Local Similarity 87.9%; Pred. No. 2.2e-281;
 Matches 1218; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
 Oy 131 TGTTCACCTTTTATCCCTGTCCTGGAAGGATATGCAAGATGAGATCTTAAGAG 190
 Db 210 TGTTCACCTTTTATCCCTGTCCTGGAAGGATATGCAAGATGAGATGTTCTACGG 265
 Oy 191 GACTTTGAGAAATTAAGAAAGAAATAGCAGCTGTGAGATGTTCTTAAGCAATCAT 250
 Db 270 AACATTTGAGAAATTAAGAAAGAAATAGCAGATTAATGAAGATGTTCTTAAGCAATCAT 329
 Oy 251 CAACCTAGCTTTTATATGTAAGCCCAAGACATCTATGAGGATGTTGCACTTCTGCT 310
 Db 330 CAACCTAGCTTTTATATGTAAGCCCAAGACATCTATGAGGATGTTGCACTTCTGCT 389
 Oy 311 TGATACCTGTTGAGCCGACAGCTGAGTGGCTCCAGAACTTAAGAAAGCAATTCAT 370
 Db 390 TGATACCTGTTGAGCCGACAGCTGAGTGGCTCCAGAACTTAAGAAAGCAATTCAT 449
 Oy 371 GTACAAAAACCTGACAGCAAGATGAGGCTGAGAAAGTTCACTGAGGCACTGAGAAATAC 430
 Db 450 GTACAAAAACCTGACAGCAAGATGAGGCTGAGAAAGTTCACTGAGGCACTGAGAAATAC 509
 Oy 431 CCATGAGGAGAGGAGAGAAATATCAGTGTGATGCTGAGGCAAGATTTCAATAGATAC 490
 Db 510 CCATGAGGAGAGGAGAGAAATATCAGTGTGATGCTGAGGCAAGATTTCAATAGATAC 569
 Oy 491 CATCTGGGCTTGTGGCAGCAGATGGGACTGCTCCAGAAAGGCAATTCACAGAAAGTTCT 550
 Db 570 TATTTAGGCTTGTGGCAGCAGATGGGACTGCTCCAGAAAGGCAATTCACAGAAAGTTCT 629
 Oy 551 GGTGGTACCTCTTTGATGATGATGCAAGAGAGGCTTCAAGCAAGAGGAAATGTT 610
 Db 630 GGTGGTACCTCTTTGATGATGATGCAAGAGAGGCTTCAAGCAAGAGGAAATGTT 689
 Oy 611 TGTTTATACCAACCTTACATCAATCTCAAGAGAGGCTGATACCGAAGCGAGGGGCG 670
 Db 690 TGTTTATACCAACCTTACATCAATCTCAAGAGAGGCTGATACCGAAGCGAGGGGCG 749
 Oy 671 GGTGGAAGCTGCAAGAGTGGAGGCTTGTGATGATGATGATGATGATGATGATGATGAT 720
 Db 750 TGTGGAAGCTGCAAGAGTGGAGGCTTGTGATGATGATGATGATGATGATGATGATGAT 809
 Oy 731 CATCTACAGTCTTCAACAGATATTCAGAAATACAGATGAGGCTGCAAGATTTCCAAAC 790
 Db 810 CATCTACAGTCTTCAACAGATATTCAGAAATACAGATGAGGCTGCAAGATTTCCAAAC 869
 Oy 791 AGCTGTATTAAGGATGAGATGCAAGATGATGATGATGATGATGATGATGATGATGAT 850
 Db 870 AGCTGTATTAAGGATGAGATGCAAGATGATGATGATGATGATGATGATGATGATGAT 929
 Oy 851 AATTGTCAATCAGCTAAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 910
 Db 930 AATTGTCAATCAGCTAAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 989
 Oy 911 TGTAGCAGATCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 970
 Db 990 TGTAGCAGATCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1049
 Oy 971 GAGCAGCTGGAGATGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1030
 Db 1050 GAGCAGCTGGAGATGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1109
 Oy 1031 GGAAGCAGCTGCTATTAAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1090
 Db 1110 GGAAGCAGCTGCTATTAAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1169
 Oy 1091 GCTCTGAGCTGAGAGAAACCAAGTGAAGTGTGCTTCCAGTATTAATCAATTACAA 1150

Dh 1170 GCTTCGACCTGACAGAGAAACAGAGAAATTGGTCCCTCCAGTAATTATGACTACATPAA 1229
Qy 1151 GGTAAATATTTTCCAACTACAGTCTGTGTAGTGAAGTCTGACGACGAAACCTTCTACCCAC 1210
Dh 1230 GGAAATATTTTCCAAAGTACAGCTGTGTATGGAAGCTGACACGAAACCTTCTTACCCAC 1289
Qy 1211 TGGGCTGCAATTCATGCGACGTGAAAAGCCAGGCGCATCATGAGAGAGGTTATGAGCT 1270
Dh 1290 TGGACTGACAGTTTCCAGGAGTGAACAAGGCGAGGCTATCATGAGAAGAGTCAATGATCT 1349
Qy 1271 GGTGCAACCCCTCAATATCATCTCAGGTCCTGAGCATGAGAGAAAGGACAGACATCACTT 1330
Dh 1350 TGTGCAACCCCTCAATATCATCTCAGGTCCTGAGCATGAGAGAAAGGACAGACATCACTT 1409
Qy 1331 TTTGATCCAGCTGAGTGTGCTGAGGACAGTCTGATGATGATGATGATGATGATGATGATG 1390
Dh 1410 CTGATTCACAGCTGAGTGTGCTGAGGACAGTCTGATGATGATGATGATGATGATGATG 1469
Qy 1391 CTTCATCATCTCCACGAGACACACATGATGATGATGATGATGATGATGATGATGATGATG 1450
Dh 1470 TTTTCATCATCTCCACGAGACACACATGATGATGATGATGATGATGATGATGATGATGATG 1529
Qy 1451 TGTGCTGCTTTGGCTGTGTGTCTTATGTTGTGTCAGACATGAGAAATGCTGCTAG 1510
Dh 1530 TGTGCTGCTGT 1589
Qy 1511 GTCCCTA 1516
Dh 1590 GTCCCTA 1595

RESULT 14
AF009513 1716 bp mRNA linear ROD 01-JUN-1998
LOCUS AF009513
DEFINITION Mus musculus hematopoietic lineage switch 2 (HLS2) mRNA, complete

ACCESSION AF009513
VERSION AF009513.1 GI:3169728
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE
AUTHORS Williams, J.H., Chan, C.-Y. and Klincken, S.P.
TITLE Hematopoietic lineage Switch 2 (HLS2), a novel mRNA species induced during an erythroid to myeloid lineage switch
JOURNAL Unpublished
2 (bases 1 to 1716)
REFERENCE
AUTHORS Williams, J.H., Chan, C.-Y. and Klincken, S.P.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-1997) Biochemistry Department (Laboratory of Cancer Medicine), University of Western Australia, Royal Perth Hospital, Wellington Street, Perth, WA 6001, Australia

FEATURES
source
1. 1716
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1. 1716
/gene="HLS2"
97. 1398
/gene="HLS2"
/codon_start=1
/product="hematopoietic lineage switch 2"
/protein_id="AAC17945.1"
/translation="MRSLEFLFIHLALSGKAVPKNVGSORTREIKELIANYEDV AKATINLAVGKYNRSYERGLVMDTPRLSGSKLKERLQIMYONLQODGLENNH LEQVRIHMERGESAVMLPRHMKMLILGSSIGTPRGSTIRLVYVAFDELQOR ASEARKIIVNPTGYTEKTVQVRQGVAAAVGAVASILQSVAFSISPTGLO KYQGVKPIPTACTIVEDAEMMSRMASGNKIVHLEMGATYPTDSFNTVAITYS MYPBDVGALDDGGAFLISWEALSLVMDLIRPKRTLRVLTMTAEKGGIGASQYVE

gene
Dh 1170 GCTTCGACCTGACAGAGAAACAGAGAAATTGGTCCCTCCAGTAATTATGACTACATPAA 1229
Qy 1151 GGTAAATATTTTCCAACTACAGTCTGTGTAGTGAAGTCTGACGACGAAACCTTCTACCCAC 1210
Dh 1230 GGAAATATTTTCCAAAGTACAGCTGTGTATGGAAGCTGACACGAAACCTTCTTACCCAC 1289
Qy 1211 TGGGCTGCAATTCATGCGACGTGAAAAGCCAGGCGCATCATGAGAGAGGTTATGAGCT 1270
Dh 1290 TGGACTGACAGTTTCCAGGAGTGAACAAGGCGAGGCTATCATGAGAAGAGTCAATGATCT 1349
Qy 1271 GGTGCAACCCCTCAATATCATCTCAGGTCCTGAGCATGAGAGAAAGGACAGACATCACTT 1330
Dh 1350 TGTGCAACCCCTCAATATCATCTCAGGTCCTGAGCATGAGAGAAAGGACAGACATCACTT 1409
Qy 1331 TTTGATCCAGCTGAGTGTGCTGAGGACAGTCTGATGATGATGATGATGATGATGATGATG 1390
Dh 1410 CTGATTCACAGCTGAGTGTGCTGAGGACAGTCTGATGATGATGATGATGATGATGATG 1469
Qy 1391 CTTCATCATCTCCACGAGACACACATGATGATGATGATGATGATGATGATGATGATGATG 1450
Dh 1470 TTTTCATCATCTCCACGAGACACACATGATGATGATGATGATGATGATGATGATGATGATG 1529
Qy 1451 TGTGCTGCTTTGGCTGTGTGTCTTATGTTGTGTCAGACATGAGAAATGCTGCTAG 1510
Dh 1530 TGTGCTGCTGT 1589
Qy 1511 GTCCCTA 1516
Dh 1590 GTCCCTA 1595

BASE COUNT 476 a 370 c 414 g 456 t
ORIGIN
LHKANISKYSLVNEADSGTFLPTGLQFTGSDKARAIKREVNLLQPLANTKVPNSBEG
TDFNFMIOAGVNPASLRDDLKYKFFHSHSDMTWSIOER"

Query Match 58.2%; Score 1077.8; DB 10; Length 1716;
Best Local Similarity 82.2%; Pred. No. 9,4e-211;
Matches 1290; Conservative 0; Mismatches 239; Indels 40; Gaps 3;

Qy 131 TGTTCACCTTTTATCCCTGTGCTGTGGAAGCAATATGAGAAGTGCATCTCTAGAG 190
Dh 123 TGTTCACCTTTTATGCTTGGCTCTGTGGAAGCAATATGAGAAGTGTCTTCTACGC 182
Qy 191 GACTTTGAGAAATATAAAGAAATAGCCAGCTGTGAGATGTTGCTAAAGCAATCAT 250
Dh 183 AACATTTGAGAAATATAAAGAAATAGCCAGCTGTGAGATGTTGCTAAAGCAATCAT 242
Qy 251 CAACCTGCTGTTTATGTTAAGCCAGAAACAGATCTATGAGGATGTCATCTTGTGT 310
Dh 243 CAACCTGCTGTTTATGTTAAGCCAGAAACAGATCTATGAGGATGTCATCTTGTGTGT 302
Qy 311 TGTACTGTTGAGCCAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCT 370
Dh 303 TGTACTGTTGAGCCAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCT 362
Qy 371 GTACCAAAACCTGAGCAAGATGAGGCTGAGAAAGTTCACTGAGGACAGTGAATACC 430
Dh 363 GTACCAAAACCTGAGCAAGATGAGGCTGAGAAAGTTCACTGAGGACAGTGAATACC 422
Qy 431 CCATGAGAGAGGAGAGAAATACGCTGTGATGCTGAGGACCAAGAAATTCATTAAGATAGC 490
Dh 423 CCATGAGAGAGGAGAGAAATACGCTGTGATGCTGAGGACCAAGAAATTCATTAAGATAGC 482
Qy 491 CATCTGCTGCTTGTGCGACGACATGAGGATCTCTTCCAGAAAGGCTTACAGCAAGATGCT 550
Dh 483 TATCTGAGTGTGCGACGACATGAGGATCTCTTCCAGAAAGGCTTACAGCAAGATGCT 542
Qy 551 GGATGACCTCTTTCATGATGAACCTGACAGAGAGGCTCTGAGAGCAAGAGGAAATGCT 610
Dh 543 GGATGACCTCTTTCATGATGAACCTGACAGAGAGGCTCTGAGAGCAAGAGGAAATGCT 602
Qy 611 TGTATTATACCAACCTTACATCACTGATCAAGAGAGGCTCAATACGAGAGGAGGCT 670
Dh 603 TGTATTATACCAACCTTACATCACTGATCAAGAGAGGCTCAATACGAGAGGAGGCT 662
Qy 671 GGTGGAAGCTGCCAAGGTGGGGCTTTGGCATCTCTATTCGATCCGTGCTCTCTCTC 730
Dh 663 TGTGGAAGCTGCCAAGGTGGGGCTTTGGCATCTCTATCCATCAATCAGTAGCTCTCTCTC 722
Qy 731 CATCTACAGTCTTACACAGGATTTTCAAGAAATCCAGATGGGGTCCCAAAATTCACAC 790
Dh 723 CATCTACAGTCTTACACAGGATTTTCAAGAAATCCAGATGGGGTCCCAAAATTCACAC 782
Qy 791 AGCTGTATTACCGGTGAGATGAGAAATGATGTCAAGATGAGCTTCTCATGAGATCAA 850
Dh 783 AGCTGTATTACCGGTGAGATGAGAAATGATGTCTCGAATGAGCTTCTCATGAGATCAA 842
Qy 851 AATTGTATTACGCTAAAGATGGGGCAAAAGACTTACCCAGATCTGATTCCTTCAACAC 910
Dh 843 AATTGTATTACGCTAAAGATGGGGCAAAAGACTTACCCAGATCTGATTCCTTCAATAC 902
Qy 911 TGTAGCAGAGATCACTGAGGAGCAAAATATCCAGAAACAGTGTGATGTCAGTGCATCT 970
Dh 903 TGTAGCAGAGATCACTGAGGAGCAAAATATCCAGAAACAGTGTGATGTCAGTGCATCT 935
Qy 971 GGAACAGCTGGAGATGTTGGGAGAGGTGTCATGATGATGAGGAGCTTTATATATCATG 1030
Dh 936 GGAACAGCTGGAGATGTTGGGAGAGGTGTCATGATGATGAGGAGCTTTATATATCATG 986
Qy 1031 GGAACAGCTGGAGATGTTGGGAGAGGTGTCATGATGATGAGGAGCTTTGCGGCTGGT 1090
Dh 987 GGAACAGCTGGAGATGTTGGGAGAGGTGTCATGATGATGAGGAGCTTTGCGGCTGGT 1046

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 12:51:18 ; Search time 136.658 Seconds
(without alignments)
11200.072 Million cell updates/sec

Title: US-09-745-763-35_COPY_57_623

Perfect score: 567
Sequence: 1 CCTATCAGATTATCTTAACA.....AGATTGTGTATTAACCAA 567

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq_19Jun03.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	567	100.0	1851	19	AAV82779
2	567	100.0	1851	21	AAZ98034
3	567	100.0	1863	21	AAZ98034
4	567	100.0	1863	22	AAZ11647
5	567	100.0	1863	24	ABK68743
6	567	100.0	1863	25	ACC50817
7	567	100.0	1863	25	ABZ71453
8	567	100.0	1884	21	AAZ58313

9	567	100.0	1895	22	AAH99703
10	567	100.0	1923	21	AAZ98139
11	567	100.0	2077	25	ACC50603
12	567	100.0	2077	25	ABZ71331
13	563.8	99.4	1860	22	AAK94491
14	555	97.9	609	20	AAV87487
15	546.4	96.4	895	21	AAH40439
16	526.4	92.8	1767	21	AAH40439
17	526.2	92.8	1778	18	AAV02296
18	525	92.6	1784	23	AAH73592
19	512.2	90.3	642	22	AAK92181
20	512.2	90.3	642	22	AAK93447
21	399.6	70.5	1778	24	ABK62598
22	290.8	51.3	314	20	AAV86369
23	284.6	50.2	462	18	AAV73398
24	284.6	50.2	462	18	AAH88060
25	284.6	50.2	462	18	AAV02139
26	277.2	48.9	441	22	AAK56831
27	254.6	44.9	317	20	AAH40587
28	121	21.3	217	20	AAH40583
29	60	10.6	60	24	ABN36212
30	49	8.6	424	23	AAH73591
31	41	7.2	65	24	ABN28454
32	39.4	6.9	975	21	AAH77173
33	39.4	6.9	1556	24	AAH3726
34	39.4	6.9	2319	25	ABK63579
35	39.4	6.9	2761	25	ABK34826
36	39.4	6.9	3383	23	ABV30133
37	39.4	6.9	8211	22	AAH80450
38	37.8	6.7	1353	24	ABN70123
39	36.8	6.5	8496	24	ABQ76621
40	36.6	6.5	766431	24	ABQ76621
41	36.2	6.4	1767	22	AAH06993
42	36	6.3	1767	24	ABJ33014
43	36	6.3	22655	22	AAH70122
44	36	6.3	2365589	24	ABH90521
45	35.4	6.2	605	22	AAH85262

ALIGNMENTS

AAV82779	standard; cDNA; 1851 BP.
AAV82779;	
25-FEB-1999	(first entry)
DE	Clone bu45_2 isolated from human adult placenta cDNA library.
XX	Secreted protein; nutritional activity; immune stimulating; vaccine;
XX	suppressing activity; haematopoiesis regulating activity;
XX	tissue growth activity; activin; chemotaxis;
KW	chemokine activity; haemostasis; thrombolytic activity; receptor;
KW	ligand; anti-inflammatory; cadherin; tumour invasion suppressor;
KW	tumour inhibition; gene therapy; ds.
XX	
OS	Homo sapiens.
XX	
PN	W09842739-A2.
XX	
PD	01-OCT-1998.
XX	
PF	20-MAR-1998; 98WO-US05653.
XX	
PR	19-MAR-1998; 98US-0044466.
XX	
PR	21-MAR-1997; 97US-0822167.
XX	
PA	(GENY) GENETICS INST INC.
XX	
PI	Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding V, Treacy M;
 XX WPI: 1998-609890/51.
 DR P-PSDB; AAM85456.
 XX
 PT New polynucleotides encoding secreted human proteins - derived from
 PT human foetal brain, adult brain, foetal kidney, placenta or adult
 PT pineal gland cDNA libraries.
 XX
 PS Claim 14; Page 69-70; 113pp; English.
 XX
 CC The present sequence encodes a secreted protein. The polynucleotide and
 CC secreted protein are predicted to have biological activities which would
 CC make them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is given.
 CC Suggested activities include nutritional activity, immune stimulating
 CC (e.g. as vaccine) or suppressing activity, haemostatic activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cachectin/tumour
 CC invasion suppressor activity, and tumour inhibition activity (no data is
 CC given in the specification to support these activities). The
 CC polynucleotide is also stated to be useful for gene therapy.
 CC
 SQ Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;
 XX
 Query Match 100.0%; Score 567; DB 19; Length 1851;
 Best Local Similarity 100.0%; Pred. No. 4,2e-161;
 Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCTATCAGATTATCTTACAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60
 DB 57 CCTATCAGATTATCTTACAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 116
 QY 61 GCATTTTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
 DB 117 GCATTTTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 176
 QY 121 GGCATCTCTAGAGACTTTGAAAGAAATTAAGAAATAGCAGCTGGAGATGTT 180
 DB 177 GGCATCTCTAGAGACTTTGAAAGAAATTAAGAAATAGCAGCTGGAGATGTT 236
 QY 181 GCTAAAGCAATCATCAACTAGCTGTTTATGTTAAAGCCAGAAACAGATCCTATGAGCA 240
 DB 237 GCTAAAGCAATCATCAACTAGCTGTTTATGTTAAAGCCAGAAACAGATCCTATGAGCA 296
 QY 241 TTGGCACTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 300
 DB 297 TTGGCACTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 356
 QY 301 GCCATCCAAATATGTAACCAAACTGAGCAAGTGGGCTGGAGAAAGTTCACTGGAG 360
 DB 357 GCCATCCAAATATGTAACCAAACTGAGCAAGTGGGCTGGAGAAAGTTCACTGGAG 416
 QY 361 CCAAGTGAATATACCCCACTGGAGAGGGGAGAAAGATCAGTGTGATCTGAGCCAAAGA 420
 DB 417 CCAAGTGAATATACCCCACTGGAGAGGGGAGAAAGATCAGTGTGATCTGAGCCAAAGA 476
 QY 421 ATTATATAGATAGCATCTGGGTTCTGGAGCAGCATTTGGGATCTCTCCAGAAAGCAT 480
 DB 477 ATTATATAGATAGCATCTGGGTTCTGGAGCAGCATTTGGGATCTCTCCAGAAAGCAT 536
 QY 481 ACACAGAAAGTTCTGGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 540
 DB 537 ACACAGAAAGTTCTGGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 596
 QY 541 AGAGGAAGATTTGTTTATACCAA 567
 DB 597 AGAGGAAGATTTGTTTATACCAA 623

ID ABQ92016 standard; cDNA; 1851 BP.
 XX
 AC ABQ92016;
 XX
 DT 04-OCT-2002 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 13.
 XX
 KW Human; cytostatic; antitubercular; antiarthritic; vulnery; analgesic;
 KW antihistaminic; antibacterial; immunosuppressive; antiparkinsonian;
 KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
 KW antidiabetic; lungicidal; antidiabetic; antidiabetic; antiallergic;
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KW stem cell; growth factor; nervous system disease; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;
 KW multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002065394-A1.
 XX
 PD 30-MAY-2002.
 XX
 PF 22-DEC-2000; 2000US-0745763.
 XX
 PR 18-MAR-1998; 98US-0040963.
 XX
 PA (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREAC/) TREACY M.
 PA (SPAU/) SPAULDING V.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 XX
 DR WPI: 2002-582343/62.
 DR P-PSDB; ABP61800.
 XX
 PT Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis
 PT
 PS Claim 50; Page 113-114; 284pp; English.
 XX
 CC The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (cDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (II) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haemopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or

CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polynucleotide of the invention.

XX
 SQ Sequence 1851 BP, 531 A, 413 C, 438 G, 469 T, 0 other;

Query Match 100.0%; Score 567; DB 24; Length 1851;

Best Local Similarity 100.0%; Pred. No. 4.2e-161; Mismatches 0; Indels 0; Gaps 0;

Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGATTCGATTATCTTAACAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60
 DB 57 CCGATTCGATTATCTTAACAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 116
 QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGAGAAAT 120
 DB 117 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGAGAAAT 176
 QY 121 GGCATCTCTAAGAGAGACTTTTGAAGAAATAAAGAGAAATAGCAGCTGTGAGATGTT 180
 DB 177 GGCATCTCTAAGAGAGACTTTTGAAGAAATAAAGAGAAATAGCAGCTGTGAGATGTT 236
 QY 181 GCTTAAGCAATCATCAACCTTACCTGTTTATGTTAAGCCAGAAACAGATCCTATGAGCA 240
 DB 237 GCTTAAGCAATCATCAACCTTACCTGTTTATGTTAAGCCAGAAACAGATCCTATGAGCA 296
 QY 241 TTGGACATCTTCTGTGTGATCTTGTGACCCGACACTGATGCTCCAGAACTTGAAGAAA 300
 DB 297 TTGGACATCTTCTGTGTGATCTTGTGACCCGACACTGATGCTCCAGAACTTGAAGAAA 356
 QY 301 GCCATCCAAATATATGACCAAAACCTGACAGAGAGTGGCTGAGAAAGTTCACTGGAG 360
 DB 357 GCCATCCAAATATATGACCAAAACCTGACAGAGAGTGGCTGAGAAAGTTCACTGGAG 416
 QY 361 CCAGTGAAGATATCCCACTGGGAGAGGAGAGAAATCACTGTGATCTGAGGCCAAGA 420
 DB 417 CCAGTGAAGATATCCCACTGGGAGAGGAGAGAAATCACTGTGATCTGAGGCCAAGA 476
 QY 421 ATTCTATAGATAGCCATCTCTGGCTTGGCAGCAGCATTTGGGACTCTCCCGAAGGCATT 480
 DB 477 ATTCTATAGATAGCCATCTCTGGCTTGGCAGCAGCATTTGGGACTCTCCCGAAGGCATT 536
 QY 481 ACAGCAGAGATTCTGTGGTGTGACCTCTTGTGATGAACTGACAGAGAGGCTCCAGAAACA 540
 DB 537 ACAGCAGAGATTCTGTGGTGTGACCTCTTGTGATGAACTGACAGAGAGGCTCCAGAAACA 596
 QY 541 AGAGGAGAGATTGTTGTTTATACCAA 567
 DB 597 AGAGGAGAGATTGTTGTTTATACCAA 623

RESULT 3

AAZ98034

ID AAZ98034 standard; cDNA; 1863 BP.

XX AAZ98034;

XX 09-MAY-2000 (first entry)

DE Human secreted protein encoding nucleotide sequence SEQ ID NO:28.

XX Human: secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiinflammatory; neurotropic; neuroprotective; antiallergic; cancer;
 KW tumour; degenerative disorder; developmental abnormality; allergy;
 KW foetal deficiency; blood disorder; immune system disorder; arthritis;
 KW autoimmune disease; hepatic disease; renal disease; inflammation;
 KW Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;

KW infection; AIDS; spinal cord injury; transplant rejection; diabetes;
 KW asthma; sepsis; acne; psoriasis; cardiovascular disorder;
 KW reproductive disorder; gastrointestinal disorder; respiratory disorder;
 KW metabolic disorder; food additive; preservative; ss.

XX Homo sapiens.

XX WO200004140-A1.

XX 27-JAN-2000.

XX 14-JUL-1999; 99WO-US15849.

XX 15-JUL-1998; 98US-0092921.

XX 15-JUL-1998; 98US-0092922.

XX 15-JUL-1998; 98US-0092956.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;
 PI Muemeki M, Endress GA, Soppet DR;

XX WPI; 2000-161128/14.

XX P-PSDB; AAY87081.

PT New isolated human genes, useful for diagnosis and treatment of, e.g.
 PT cancers, neurological or blood disorders

XX Claim 1; Page 319; 494pp; English.

CC The polynucleotide sequences given in AAZ98017 to AAZ98108 encode the
 CC human secreted proteins given in AAY87064 to AAY8723. Human secreted
 CC protein can have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;
 CC immunosuppressive; antiinflammatory; neurotropic; neuroprotective; and
 CC antiallergic. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Human secreted proteins and their polynucleotides can
 CC be used for developing products for the diagnosis or treatment of cancer,
 CC tumours, neurodegenerative disorders, developmental abnormalities and
 CC foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation,
 CC allergies, Alzheimer's disease, behavioural disorder, schizophrenia,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The
 CC proteins or polynucleotides can also be used as food additives or
 CC preservatives. The proteins are also useful for identifying their
 CC binding partners. AAZ98008 to AAZ98016 and AAY87063 are sequence used in
 CC the exemplification of the present invention.

XX Sequence 1863 BP, 533 A, 417 C, 443 G, 470 T, 0 other;

Query Match 100.0%; Score 567; DB 21; Length 1863;

Best Local Similarity 100.0%; Pred. No. 4.2e-161; Mismatches 0; Indels 0; Gaps 0;

Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGATTCGATTATCTTAACAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60
 DB 57 CCGATTCGATTATCTTAACAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 116
 QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGAGAAAT 120
 DB 117 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGAGAAAT 176
 QY 121 GGCATCTCTAAGAGAGACTTTTGAAGAAATAAAGAGAAATAGCAGCTGTGAGATGTT 180
 DB 177 GGCATCTCTAAGAGAGACTTTTGAAGAAATAAAGAGAAATAGCAGCTGTGAGATGTT 236

Db 417 CCACGTGAGAAATACCCACCTGGAGAGGGGAGAAAGATACGCTGTGATCTGTAGCCAGACAGA 476
Qy 421 ATTGATAGATAGATGACCATCTGGGCTTTGGACAGACATTTGGGACCTCCCGAGAAAGGCAAT 480
Db 477 ATTGATAGATAGATGACCATCTGGGCTTTGGACAGACATTTGGGACCTCCCGAGAAAGGCAAT 536
Qy 481 ACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCTTCAGAAAGCA 540
Db 537 ACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCTTCAGAAAGCA 596
Qy 541 AGAGGGAAGATTTGTTTATTAACCA 567
Db 597 AGAGGGAAGATTTGTTTATTAACCA 623

RESULT 5

ABK69743
ID ABK69743 standard; cDNA; 1863 BP.

ABK69743;

DT 15-JUL-2002 (first entry)

DE Human secreted protein gene 18 #1.

Human; ssi; gene; secreted protein; gene therapy; immunosuppressive;
antiarrhythmic; antithrombotic; antiproliferative; cytototoxic; cardiac;
vasoactive; cerebroprotective; neuroprotective; anti-bacterial;
viral; fungicide; antifungal; ophthalmological; autoimmune disease; neoplasm;
rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
angiogenesis; nervous system disorder; Alzheimer's disease; infection;
ocular disorder; corneal infection; wound healing; skin aging;
epithelial cell proliferation; food additive.

Homo sapiens.

WO200226931-A2.

04-APR-2002.

24-SEP-2001; 2001WO-US29871.

25-SEP-2000; 2000US-234925P.

12-JAN-2001; 2001WO-US00911.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Komatsu J, Duan DR, Rosen CA, Moore PA, Shi Y,
Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR,
Bridges GA, Mucenki M, Edner R;

WPI; 2002-362489/39.

P-PSDB; ABG33880.

Novel 71 isolated secreted polypeptides and polynucleotides encoding
the polypeptides, useful for treating Huntington's disease, sepsis,
meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
asthma

Claim 1; Page 1170; 1478pp; English.

The invention relates to an isolated nucleic acid molecule (or its
fragment, homologue complement or allelic variant) encoding a human
secreted protein (and its fragment, domain, epitope, variant, secreted
form and species variant). Also included are a recombinant vector
comprising the nucleic acid, a recombinant host cell comprising the
vector, an antibody against the secreted protein, a recombinant host cell
that expresses the secreted protein and a method of identifying a binding
partner of the secreted protein. The nucleic acid and protein are used to
prevent, diagnose, treat or ameliorate a medical condition in e.g.
humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep

CC for example autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiodysplasia, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC disorders are listed in the specification. Many other diseases and
CC disorders are listed in the specification. The polypeptides can also be
CC used to aid wound healing an epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities. The present
CC sequence encodes a novel human secreted protein of the invention.

SO Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 100.0%; Score 567; DB 24; Length 1863;
Best Local Similarity 100.0%; Pred. No. 4.2e-161;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTATGATGATATCTTAACAAAGAAACCACTGGAAGAAAGAAATCTTATCTTC 60
Db 57 CCTATGATGATATCTTAAAGAAAGAAACCACTGGAAGAAAGAAATCTTATCTTC 116
Qy 61 GCATTTTTCGGTGTGTGACCTTTTATCCCTGTCTGTGGAAAGCTATATGCAAGAT 120
Db 117 GCATTTTTCGGTGTGTGACCTTTTATCCCTGTCTGTGGAAAGCTATATGCAAGAT 176
Qy 121 GGCATCTCTTAAGAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCAGCTGTGAGATTT 180
Db 177 GGCATCTCTTAAGAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCAGCTGTGAGATTT 236
Qy 181 GCTTAAGCATCATCACTACCTGCTGTGTTATGTTAAAGCCCAAGACAGATCTTATAGCGA 240
Db 237 GCTTAAGCATCATCACTACCTGCTGTGTTATGTTAAAGCCCAAGACAGATCTTATAGCGA 296
Qy 241 TTGGCAGCTTCTGTGTGATGCTGTGACCAAGCACTGAGCTGCAAGAACTTGAAGAA 300
Db 297 TTGGCAGCTTCTGTGTGATGCTGTGACCAAGCACTGAGCTGCAAGAACTTGAAGAA 356
Qy 301 GCCATCCAAATTATGTAACCAAAACCTGACAGAGATGGCTGAGAAAGTTCACTGAG 360
Db 357 GCCATCCAAATTATGTAACCAAAACCTGACAGAGATGGCTGAGAAAGTTCACTGAG 416
Qy 361 CCAGTGAAGATATCCCACTGGAGAGAGAGAGAGAGATCACTGTGATCTGAGCCAGA 420
Db 417 CCAGTGAAGATATCCCACTGGAGAGAGAGAGAGAGATCACTGTGATCTGAGCCAGA 476
Qy 421 ATTGATAGATAGGACATCTGGGTCTTGGAGAGAGATTTGGGATCTCTCCAGAGGCAAT 480
Db 477 ATTGATAGATAGGACATCTGGGTCTTGGAGAGAGATTTGGGATCTCTCCAGAGGCAAT 536
Qy 481 ACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCTTCAGAAAGCA 540
Db 537 ACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCTTCAGAAAGCA 596
Qy 541 AGAGGGAAGATTTGTTTATTAACCA 567
Db 597 AGAGGGAAGATTTGTTTATTAACCA 623

RESULT 6

ACC50817
ID ACC50817 standard; cDNA; 1863 BP.

ACC50817;

DT 12-JUN-2003 (first entry)

DE Human secreted protein coding sequence, SEQ ID 484.

Cardiant; antiarrhythmic; antiarteriosclerotic; vasoactive; cytototoxic;
vulnerable; antiinflammatory; neuroprotective;

KM antiparkinsonian; gene therapy; human; cardiovascular disorder;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200295010-A2.
 PD 28-NOV-2002.
 XX
 PE 19-MAR-2002; 2002MO-US09785.
 XX
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-129429/12.
 XX
 PT Novel human secreted proteins, useful for detecting, preventing,
 PT diagnosing, prognosticating, treating and/or ameliorating
 PT cardiovascular disorders such as arrhythmia -
 XX
 PS Claim 21; SEQ ID 484; 1881bp; English.
 XX
 CC The present invention relates to novel human secreted proteins
 CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The
 CC proteins and their coding sequences are useful for the preparation of a
 CC diagnostic or pharmaceutical composition for diagnosing or treating a
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
 CC coronary arteriosclerosis and myocardial ischaemia), neural disorders,
 CC immune system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of haematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of
 CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism.
 CC Note: The sequenced data for this patent was published in electronic
 CC format and is available from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
 Query Match 100.0%; Score 567; DB 25; Length 1863;
 Best Local Similarity 100.0%; Pred. No. 4, 2e-161;
 Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCTATCAGATTATCTTAAACAAGAAAAACAACCTGAAAAAAATGAATTCCTTATCTTC 60
 DB 57 CCTATCAGATTATCTTAAACAAGAAAAACAACCTGAAAAAAATGAATTCCTTATCTTC 116
 QY 61 GCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGAT 120
 DB 117 GCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGAT 176
 QY 121 GGCATCTCTAAGAGAGCTTTTGAAGAATAAAGAAATATGACAGCTGGAGATGTT 180
 DB 177 GGCATCTCTAAGAGAGCTTTTGAAGAATAAAGAAATATGACAGCTGGAGATGTT 236
 QY 181 GCTAAAGCAATCATCACTAGCTGTTTATGTTAAGCCAGACAGATCTCATGAGCA 240
 DB 237 GCTAAAGCAATCATCACTAGCTGTTTATGTTAAGCCAGACAGATCTCATGAGCA 296

QY 241 TTGGCACTTCGTTGTTATACGTTTGACCCAGACTGAGTGGCTCAAGAACTTAGAAAA 300
 DB 297 TTGGCACTTCGTTGTTATACGTTTGACCCAGACTGAGTGGCTCAAGAACTTAGAAAA 356
 QY 301 GCCATCCAAATTATATGACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTTGAG 360
 DB 357 GCCATCCAAATTATATGACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTTGAG 416
 QY 361 CCACTGAGATACCCCACTGGGAGAGGGAGAGAAATACAGCTGTGATGCTGAGCCCAAGA 420
 DB 417 CCACTGAGATACCCCACTGGGAGAGGGAGAGAAATACAGCTGTGATGCTGAGCCCAAGA 476
 QY 421 ATTGATAGATAGGCAATCCTGAGCTGCTGGAGCAGACATTTGGGACTCTCCAGAAAGGCA 480
 DB 477 ATTGATAGATAGGCAATCCTGAGCTGCTGGAGCAGACATTTGGGACTCTCCAGAAAGGCA 536
 QY 481 ACAGCAGAGTTCTGATGCTGACCTCTTTCATGATGACAGAGAGGCTCAGAAACA 540
 DB 537 ACAGCAGAGTTCTGATGCTGACCTCTTTCATGATGACAGAGAGGCTCAGAAACA 596
 QY 541 AGAGGGAAGATGTTGTTTATTAACCA 567
 DB 597 AGAGGGAAGATGTTGTTTATTAACCA 623
 RESULT 7
 AB271453
 ID AB271453 standard; cDNA; 1863 BP.
 XX
 AC AB271453;
 XX
 DT 04-APR-2003 (first entry)
 XX
 DE Secreted protein-encoding gene 142 cDNA clone HRA035, SEQ ID NO:274.
 XX
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
 KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
 KW antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200276488-A1.
 PD 03-OCT-2002.
 XX
 PE 19-MAR-2002; 2002MO-US08276.
 XX
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-029900/02.
 DR P-PSDB; ABR00274.
 XX
 PT New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 PT e.g. gastrointestinal diseases and disorders, or cancers -
 XX
 PS Claim 21; Page 918; 1216bp; English.
 CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,

Db	287	GCTAAAGCATCATCAACCTAGCTGTTTATGTTAAAGCCAGAAAGATCCTATGAGCA	346
Qy	241	TTGGCACTTCTGGTTGATCTGTTGGAGCCCAAGCTGAATGGCTCCAGAACCTTGA	3000
Db	347	TTGGCACTTCTGGTTATATACTGTGGAGCCCAAGCTGAATGGCTCCAGAACCTTGA	406
Qy	301	GCCATCCAAATTATGTATACCAAAACCTGCGAGCAAGATGGGCTGGAGAAAGTTCACTGGAG	360
Db	407	GCCATCCAAATTATGTATACCAAAACCTGCGAGCAAGATGGGCTGGAGAAAGTTCACTGGAG	466
Qy	361	CCAGTGAATATCCCACTGGAGAGGGGAGAAAGATCAGCTGTGATCTGGAGCCAGA	420
Db	467	CCAGTGAATATCCCACTGGAGAGGGGAGAAAGATCAGCTGTGATCTGGAGCCAGA	520
Qy	421	ATTCAATAGATAGCCATCTCTGGGCTTTGGAGACAGCATTTGGAGCTCTCCAGAAAGCAT	480
Db	527	ATTCAATAGATAGCCATCTCTGGGCTTTGGAGACAGCATTTGGAGCTCTCCAGAAAGCAT	586
Qy	481	ACAGCAGAAATTTCTGGTGGTGAATCTTTCCATGAACTGCAAGAAAGGGCTCAGAAAGCA	540
Db	587	ACAGCAGAAATTTCTGGTGGTGAATCTTTCCATGAACTGCAAGAAAGGGCTCAGAAAGCA	646
Qy	541	AGAGGGAAGATTTGTTTATTAACCA	567
Db	647	AGAGGGAAGATTTGTTTATTAACCA	673

RESULT 9
AAH99703
ID AAH99703 standard; cDNA; 1895 BP.

АС ААН99703;

DT 16-OCT-2001 (first entry)

DE Human protein encoding cDNA sequence SEQ ID NO:538.

KM Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KM antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KM antibacterial; endocrine; cardiac; central nervous system; vitruide
KM anti-HIV; fungicide; antimetange; cardiovascular; antihaemic; anaemic
KM antiaggregant; haemostatic; vulnary; antulcer; osteopathic; asthma;
KM demagogical; antiallergic; antiaesthetic; antiabiotic; cytostatic;
KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infect
KM immunostimulant; gene therapy; antiense therapy; vaccine; inflamm
KM antiinflammatory; rheumatoid arthritis; septic shock; pancreatitis;
KM cardiac dysfunction; neuropathologic; cardiac anaphylaxis; autotransmi
KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
KM thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KM allergic rhinitis; diabetes; multiple sclerosis; depression;
KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder
KM neurological disorder; ss.

OS Homo sapiens.

PN W0200153455-A2.

PD 26-JUL-2001

PF 22-DEC-2000; 2000WO-US35017.

PR 23-DEC-1999; 99US-0471275.

PR 25-APR-2000; 2000US-0552317.

PA (HYSE-) HYSEQ INC

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-457603/49.

XX

PT Isolated human polynucleotides encoding polypeptides, useful for the

PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS Claim 1; Page 591; 1217pp; English.
vz

CC AAH91916 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antiinflammatory; antihemetic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; virocidic; anti-HIV; fungicidal; antimutagen;
CC cardiovascular; antianemic; antiaggregant; haemostatic; vulnery;
CC anticancer; osteopathic; dermatological; antiallergic; antaesthetic;
CC antidiabetic; cytosstatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmune, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.

SQ Sequence 1895 BP; 530 A; 439 C; 450 G; 476 T; 0 other.

Query Match	100.0%	Score 567;	DB 22;	Length 1895;
Best Local Similarity	100.0%	Pred. No. 4.2e-16;		
Matches 567; Conservative	0;	Mismatches	0;	Gaps 0;

QY 1 CTTATCAGATTATCTTAACAAGAAACCCTGGAATAAATGAATTCTTATCTTC 60

Db 84 CCTATCAGATTATCTTAACAAGAAAACCAACTGGAAAAAAATGAAATTCTTATCTTC 143

61 GCATTTTCGGTGTTCACCTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGAT 120

Db 144 GCATTTTCGGTGTTCACCTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGAT 203

121 GGCATCTCTAGAGGACTTTGAAGAAATAAAGAAGAAATAGCCAGCTGTGAGATGTT 180

Db 204 GGCATCTCTAAGAGGACTTTTGAAGAAATAAAGAGAATAGCCAGCTGTGAGATGTT 263

181 GCTAAGCAATCATCAACCTAGCTGTTATGTGTAAGCCAGACAGATCCTATGAGCGA 240

Db 264 GCTAAGCAATCATCAACCTAGCTGTTATGTGTAAGCCAGACAGATCCTATGAGCGA 323

241 TTGGCACTTCTGGTTGATACTGTGGACCCAGACTGAGTGGCTCCAAGAACCTAGAAAA 300

Db 324 TTGGCACTTCTGGTTGATACTGTGGACCCAGACTGAGTGCTCCAAGAACTAGAAAA 383

301 GCCATCCAATTATGTACCAAACTGCAGCAAGATGGGCTGGAGAACTCACCCTGAG 360

Db 384 GCCATCCAATTATGTACCAAACTGCAGCAAGATGGGCTGGAGAAATTACCTGGAG 443

361 CCAGTGAGATACCCCACTGGAGAGGGGAGAAGATCAGCTGTGATGCTGGAGCCAAGA 420

Db 444 CCAGTGAGATACCCCACTGGAGAGGGGAGAGAATCAGCTGTGATGCTGGAGCCAGA 503

421 ATTCATAGATAGCCATCCTGGGCTTTGGCAGCAGCATTTGGACTCCTCCAGAAGGCATT 480

Db 504 ATTCTATAAGATAGCCATCTGGTCTTGGCAGCAGCATTGGGACTCTCCAGAAGGCATT 563

481 ACAGCAGAAGTTCTGGTGTCACCTCTTCGATGAACTGCAGAGAAGGCCCTCAGAAGCA 540

Db 564 ACAGCAGAATTCTGTGTGACCTCTTCGATGA CTGCAGAGAAGGCCCTCAGAAGCA 623

QY 541 AGAGGGAAGATTGTTGTTTATAACCA 567

Db 624 AGAGGAGAGATTGTTGTTTATAACCA 650

RESULT 10
AAZ98139
ID AAZ98139 standard; cDNA; 1923 BP.
XX
AC AAZ98139;
XX
DT 11-MAY-2000 (first entry)
XX
DE Human signal peptide containing protein HSPB-31 cDNA SEQ ID NO:165.
XX
KW Human, signal peptide-containing protein; HSPB; diagnosis; cancer;
KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;
KW antimicrobial; neuroprotective; cardiovascular; hepatotropic;
KW antiasthmatic; gene therapy; cell proliferation; neurological disorder;
KW reproductive disorder; developmental disorder; arteriosclerosis;
KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
KW Parkinson's disease; Huntington's disease; ovulatory defect;
KW muscular dystrophy; ss.
XX
OS Homo sapiens.
XX
PN W0200000610-A2.
XX
PD 06-JAN-2000.
XX
PF 25-JUN-1999; 99MO-US14484.
XX
PR 26-JUN-1998; 98US-0090762.
PR 31-JUL-1998; 98US-0094983.
PR 01-OCT-1998; 98US-0102686.
PR 11-DEC-1998; 98US-0112129.
XX
XX (INCY-) INCYTE PHARM INC.
XX
PI Lei P, Tang YT, Gorgone GA, Corley NC, Guegler KU, Baughn MR;
PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
PI Bandman O;
XX
DR WPI: 2000-160673/14.
DR P-PSDB; AAY87254.
XX
PT New human signal peptide-containing proteins useful in treatment,
PT prevention and diagnosis of e.g. cancer, inflammation and
PT cardiovascular disease -
XX
PS Claim 9; Page 269-270; 327pp; English.
XX
XX AAZ98109 to AAZ98242 encode AAY87224 to AAY87357 which represent the
XX human signal peptide-containing proteins HSPB-1 to HSPB-134. HSPBs have
XX anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,
XX neuroprotective, cardiovascular and antiasthmatic activities, and can
XX be used in gene therapy. HSPBs can be used to treat or prevent disorders
XX associated with decreased activity or function of HSPB. Antagonists of
XX HSPB are used to treat or prevent disorders associated with increased
XX activity or function of HSPB. Such diseases include cell proliferation
XX (including cancer), inflammation, cardiovascular, neurological,
XX reproductive or developmental disorders, (e.g. arteriosclerosis,
XX cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
XX asthma, Crohn's disease, microbial or other infections, congestive or
XX ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
XX diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPB
XX nucleic acids can be used for the recombinant production of HSPB, for
XX detecting HSPB in standard hybridisation and amplification assays (for
XX diagnosis and monitoring), in gene therapy, as antisense,
XX triplex-forming or ribozyme therapeutics, for detecting related sequences
XX or genetic variations, and for chromosomal mapping. HSPB are also used to
XX raise specific antibodies (Ab) and to screen for agonists and
XX antagonists (potential therapeutic agents). Ab are used to diagnose, or
XX monitor, HSPB-related diseases (in usual immunoassays), as therapeutic
XX antagonists, in competitive drug screens, and for purification of HSPB
XX from natural sources.

SQ Sequence 1923 BP; 538 A; 439 C; 471 G; 475 T; 0 other;
Query Match 100.0%; Score 567; DB 21; Length 1923;
Best Local Similarity 100.0%; Pred. No. 4.3e-161;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTATCAGATTATCTTAACAAGAAAGAAACCAACTGGAAAAAATGAAATTCCTTATCTTC 60
DB 137 CCTATCAGATTATCTTAACAAGAAAGAAACCAACTGGAAAAAATGAAATTCCTTATCTTC 196
QY 61 GCATTTTCGGTGTGTTTACCTTTATCCCTGTCTGTGGAAAGCTATATGCAAGAT 120
DB 197 GCATTTTCGGTGTGTTTACCTTTATCCCTGTCTGTGGAAAGCTATATGCAAGAT 256
QY 121 GGCACTCTTAAGAGGACTTTTGAAGAAATAAAGAAATAGCAGCTGGAGATGT 180
DB 257 GGCACTCTTAAGAGGACTTTTGAAGAAATAAAGAAATAGCAGCTGGAGATGT 316
QY 181 GCTAAAGCAATCATCACTAGCTGTTTATGTAAAGCCAGAAACAGATCCTATGACGA 240
DB 317 GCTAAAGCAATCATCACTAGCTGTTTATGTAAAGCCAGAAACAGATCCTATGACGA 376
QY 241 TTGGCACTTCTGTTGATPACTGTTTGAACCAAGCTGATGCTGCAAGAACTAGAAAA 300
DB 377 TTGGCACTTCTGTTGATPACTGTTTGAACCAAGCTGATGCTGCAAGAACTAGAAAA 436
QY 301 GCCATCCAAATTAATGTCACCAAAACCTGCACCAAGATGGGCTGGAAGAAATTCACCTGGAG 360
DB 437 GCCATCCAAATTAATGTCACCAAAACCTGCACCAAGATGGGCTGGAAGAAATTCACCTGGAG 496
QY 361 CCAATGAGAAATACCCCACTGGAGAGAGGAGAAAGAAATCAGCTGATGCTGAGCCAAAGA 420
DB 497 CCAATGAGAAATACCCCACTGGAGAGAGGAGAAAGAAATCAGCTGATGCTGAGCCAAAGA 556
QY 421 ATTCAATAGATAGGCACTCTGGGTCTTGGCAAGCAAGATGGGACTCTCCAGAAAGCAATT 480
DB 557 ATTCAATAGATAGGCACTCTGGGTCTTGGCAAGCAAGATGGGACTCTCCAGAAAGCAATT 616
QY 481 ACAGCAGAAGCTCGGTGGTGAACCTCTTGGATGAACCTGCAAGAAAGGCTCGAAGCA 540
DB 617 ACAGCAGAAGCTCGGTGGTGAACCTCTTGGATGAACCTGCAAGAAAGGCTCGAAGCA 676
QY 541 AGAGGAAAGATTGTTGTTTAAACCAA 567
DB 677 AGAGGAAAGATTGTTGTTTAAACCAA 703
RESULT 11
ACCS0603
ID ACCS0603 standard; cDNA; 2077 BP.
XX
AC ACCS0603;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted protein coding sequence, SEQ ID 270.
XX
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;
KW vulnerrary; antiinflammatory; neurotropic; neuroprotective;
KW antiparkinsonian; gene therapy; human; cardiovascular disorder;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN W0200295010-A2.
XX
PD 28-NOV-2002.
XX
PF 19-MAR-2002; 2002MO-US09785.
XX
PR 21-MAR-2001; 2001US-277340P.
PR 19-JUL-2001; 2001US-306171P.
PR 13-NOV-2001; 2001US-331287P.

XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM;
 PI WPI, 2003-129429/12.
 XX Novel human secreted proteins, useful for detecting, preventing,
 PT diagnosing, prognosticating, treating and/or ameliorating
 PT cardiovascular disorders such as arrhythmia -
 XX Claim 21, SEQ ID 270, 1881pp; English.
 XX The present invention relates to novel human secreted proteins
 CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50866). The
 CC proteins and their coding sequences are useful for the preparation of a
 CC diagnostic or pharmaceutical composition for diagnosing or treating a
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
 CC coronary arteriosclerosis and myocardial ischaemia), neural disorders,
 CC immune system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of
 CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism.
 CC Note: The sequence data for this patent was published in electronic
 CC format and is available from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 2077 BP, 597 A, 436 C, 482 G, 562 T, 0 other;
 SQ
 Query Match 100.0%; Score 567, DB 25; Length 2077;
 Best Local Similarity 100.0%; Pred. No. 4,4e-161;
 Matches 567, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 510 ATTCTAAGATATGACCACTCGGTCTTGCGACGACGCTTGGCTCTCCAGAAAGGCAATT 569
 QY 481 ACAGCAGAGTTCTGCGGTGACCTCTTTCATGTAACCTGACGAGAAAGGCTCAGAAACA 540
 DB 570 ACAGCAGAGTTCTGCGGTGACCTCTTTCATGTAACCTGACGAGAAAGGCTCAGAAACA 629
 QY 541 AGAGGAGAGTTGTTGTTTAAACCA 567
 DB 630 AGAGGAGAGTTGTTTAAACCA 656
 RESULT 12
 ID ABZ71331 standard; cDNA; 2077 BP.
 AC ABZ71331;
 DT 04-APR-2003 (first entry)
 XX Secreted protein-encoding gene 142 cDNA clone HRACJ35, SEQ ID NO:152.
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
 KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytogenetic; gene therapy;
 KW antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2;
 KW gene; ss.
 XX Homo sapiens.
 OS MO200276488-A1.
 XX MO200276488-A1.
 XX 03-OCT-2002.
 PD 19-MAR-2002; 2002MO-US08276.
 PF 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM;
 DB WPI, 2003-029900/02.
 DR P-PSDB; ABR00152.
 XX New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating,
 PT e.g. gastrointestinal diseases and disorders, or cancers -
 XX Claim 21, Page 841-842; 1216pp; English.
 PS ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight

CC markers. The present sequence represents a human secreted protein-
XX encoding cDNA clone of the invention.

CC Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other:

Query Match 100.0%; Score 567; DB 25; Length 2077;

Best Local Similarity 100.0%; Pred. No. 4,4e-161;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCTTCAGATTATCTTAAACAAGAAACCACTGAAAAAATAATTCCTTATCTTC 60
D 90 CTTATCGATTATCTTAAACAAGAAACCACTGAAAAAATAATTCCTTATCTTC 149
QY 61 GCATTTTCGGTGTGTGTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
D 150 GCATTTTCGGTGTGTGTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 209
QY 121 GGCATCTCTAAGAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 180
D 210 GGCATCTCTAAGAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 269
QY 181 GCTAAGCAATCATCACTAGCTGTTTATAGTAAAGCCCAAGAACATCTTATGAGCGA 240
D 270 GCTAAGCAATCATCACTAGCTGTTTATAGTAAAGCCCAAGAACATCTTATGAGCGA 329
QY 241 TTGGCACTTCTGTGTGATGATGTCAGCCAGACTGATGCTCCAGAAACCTTGAAGAAA 300
D 330 TTGGCACTTCTGTGTGATGATGTCAGCCAGACTGATGCTCCAGAAACCTTGAAGAAA 389
QY 301 GCCATCCAAATTATGTACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTGGAG 360
D 390 GCCATCCAAATTATGTACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTGGAG 449
QY 361 CCAATGGAATACCCCACTGGGAGAGAGGAGAAAGATAGCTGTGATGCTGAGCCAGAA 420
D 450 CCAATGGAATACCCCACTGGGAGAGAGGAGAAAGATAGCTGTGATGCTGAGCCAGAA 509
QY 421 ATTATAGATAGGACATCTGGGTCTTTGGCAGACGATTTGGGAACTCCCTCAGAAAGGCA 480
D 510 ATTATAGATAGGACATCTGGGTCTTTGGCAGACGATTTGGGAACTCCCTCAGAAAGGCA 569
QY 481 ACACAGAGAGTCTGTGTGTGACCTCTTCGATGAATGACAGAGAGGGCTCAGAAAGCA 540
D 570 ACACAGAGAGTCTGTGTGTGACCTCTTCGATGAATGACAGAGAGGGCTCAGAAAGCA 629
QY 541 AGAGGGAAGATTGTTGTTTATACCAA 567
D 630 AGAGGGAAGATTGTTGTTTATACCAA 656
```

RESULT 13
AAK94491
ID AAK94491 standard; cDNA; 1860 BP.

XX AAK94491;

DT 06-NOV-2001 (first entry)

DE Human full-length cDNA, SEQ ID NO: 3328.

KM Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

PN EP1130094-A2.

PD 05-SEP-2001

PP 07-JUL-2000; 2000EP-0114089.

XX 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI; 2001-524255/58.

PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -

PS Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.

CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other:

Query Match 99.4%; Score 563.8; DB 22; Length 1860;
Best Local Similarity 99.6%; Pred. No. 3.9e-160;
Matches 565; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 CTTATCGATTATCTTAAACAAGAAACCACTGAAAAAATAATTCCTTATCTTC 60
D 87 CTTATCGATTATCTTAAACAAGAAACCACTGAAAAAATAATTCCTTATCTTC 146
QY 61 GCATTTTCGGTGTGTGTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
D 147 GCATTTTCGGTGTGTGTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 206
QY 121 GGCATCTCTAAGAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 180
D 207 GGCATCTCTAAGAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 266
QY 181 GCTAAGCAATCATCACTAGCTGTTTATAGTAAAGCCCAAGAACAGATCCTATGAGCGA 240
D 267 GCTAAGCAATCATCACTAGCTGTTTATAGTAAAGCCCAAGAACAGATCCTATGAGCGA 326
QY 241 TTGGCACTTCTGTGTGATGATGTCAGCCAGACTGATGCTCCAGAAACCTTGAAGAAA 300
D 327 TTGGCACTTCTGTGTGATGATGTCAGCCAGACTGATGCTCCAGAAACCTTGAAGAAA 386
QY 301 GCCATCCAAATTATGTACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTGGAG 360
D 387 GCCATCCAAATTATGTACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTGGAG 446
QY 361 CCAATGGAATACCCCACTGGGAGAGAGGAGAAAGATAGCTGTGATGCTGAGCCAGAA 420
D 447 CCAATGGAATACCCCACTGGGAGAGAGGAGAAAGATAGCTGTGATGCTGAGCCAGAA 506
QY 421 ATTATAGATAGGACATCTGGGTCTTTGGCAGACGATTTGGGAACTCCCTCAGAAAGGCA 480
D 507 ATTATAGATAGGACATCTGGGTCTTTGGCAGACGATTTGGGAACTCCCTCAGAAAGGCA 566
QY 481 ACACAGAGAGTCTGTGTGTGACCTCTTCGATGAATGACAGAGAGGGCTCAGAAAGCA 540
D 567 ACACAGAGAGTCTGTGTGTGACCTCTTCGATGAATGACAGAGAGGGCTCAGAAAGCA 626
QY 541 AGAGGGAAGATTGTTGTTTATACCAA 567
D 627 AGAGGGAAGATTGTTGTTTATACCAA 653
```

RESULT	14
AAV87487	ID AAV87487 standard; cDNA, 609 BP.
XX	
AC	AAV87487;
XX	
DT	27-APR-1999 (first entry)
XX	
DE	EST clone BU45.
XX	
KW	Expressed sequence tag; secreted protein; hematopoiesis regulator;
KW	tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW	chemokinesis; chemokinesis; haemostasis; gene therapy; thrombolysis;
XX	receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
OS	Homo sapiens.
XX	
PN	WO9845435-A2.
PD	15-OCT-1998.
XX	
PF	10-APR-1998; 98WO-US06954.
PR	10-APR-1997; 97US-0835913.
XX	(GENY) GENETICS INST INC.
PA	
PI	Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
P1	Racte LA, Spaulding V, Treacy M;
XX	WPI; 1999-070076/06.
DR	
PT	New polynucleotides encoding human secreted proteins - derived from
PT	e.g. human blood, kidney, foetal lung, placenta, testes, brain,
FT	ovary, pituitary, retina and colon cDNA libraries
XX	
PS	Claim 1; Page 580; 633pp; English.
XX	
CC	This sequence represents an expressed sequence tag (EST), and is a
CC	polynucleotide of the invention. The polynucleotides of the invention are
CC	all secreted EST sequences isolated from a variety of human tissue
CC	sources. The EST sequences and proteins encoded by them are predicted to
CC	have useful biological activities which would make them suitable for
CC	treating, preventing or ameliorating medical conditions in humans and
CC	animals, although no supporting data is given. Suggested activities
CC	include nutritional activity, immune stimulating or suppressing activity,
CC	haematopoiesis regulating activity, tissue growth activity,
CC	activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC	activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC	activity. The EST sequences are also stated to be useful for gene
CC	therapy.
XX	
QO	Sequence 609 BP; 190 A; 127 C; 148 G; 144 T; 0 other;
Query Match	97.9%; Score 555; DB 20; Length 609;
Best Local Similarity	99.8%; Pred. No. 1.1e-157;
Matches	566; Conservative 0; Mismatches 0; Indels 1; Gaps 1
OY	1 CCTATCAGATTATCTTTAACAGAAAAACCACGTAAGAAAAAATGAATTCTTATCTTC 60
Dd	29 CCTATCACATTTATCTTTAACAGAAAAACCACGTAAGAAAAAATGAATTCTTATCTTC 88
OY	61 GCATTTTCGGGCGGTTCACCTTTTATGCCCTGCTCGTGGAAGACTTATGCAGAAGAT 120
Dd	89 GCATTTTCGGGCGGTTCACCTTTTATGCCCTGCTCGTGGAAGACTTATGCAGAAGAT 148
OY	121 GGCACTCTTAAGAGAGACTTTTGAAGAAATAAAGAAAGATGCCAGCTGTGAGATGTT 180
Dd	149 GGCACTCTTAAGAGAGACTTTTGAAGAAATAAAGAAAGATGCCAGCTGTGAGATGTT 208
OY	181 GCTAAGCAATCATCAACCTGCTTTATYGSTAAAGCCAGAAACAGATCTTAGAGACA 240

Db	209	6CTAAAGCANTCATCACTTACTCTTTTATGTGTAAAGCCAGAAACAGTCTTAGACGA	268		
Qy	241	TTGGCAGCTTCTGGTTGATATCTGTTGACCCAGACTGAGTGGCTCCAGAACTGAAAAA	3000		
Db	269	TTGGCAGCTTCTGGTTGATCTGTTGACCCAGACTGAGTGGCTCCAGAA-CTGAAAAA	3270		
Qy	301	GCCATCCAAATTAATGACCAAAACCTGCAGCAAGATGGGGCTGAGAAAAGTTCACTGAG	3660		
Db	328	GCCATCCAAATTAATGACCAAAACCTGCAGCAAGATGGGGCTGAGAAAAGTTCACTGAG	3870		
Qy	361	CCAGTAGAATAATCCCACTGGAGAGAGGAGAAAGTAATAGCTGTGATCTGGAGCCAGA	4200		
Db	388	CCAGTAGAATAATCCCACTGGAGAGAGGAGAAAGTAATAGCTGTGATCTGGAGCCAGA	4440		
Qy	421	ATTCAATAGATAGCCATCTCTGGGCTTGTGACAGACATTTGGGACTCTCCAGAAAGCATT	4800		
Db	448	ATTCAATAGATAGCCATCTCTGGGCTTGTGACAGACATTTGGGACTCTCCAGAAAGCATT	5070		
Qy	481	ACAGCAGAAAGTTCTGGTGTGATCTCTTTCCATGAACTGCAGAGAAAGGGCTCAGAGCA	5400		
Db	508	ACAGCAGAAAGTTCTGGTGTGATCTCTTTCCATGAACTGCAGAGAAAGGGCTCAGAGCA	5670		
Qy	541	AGAGGAGAGATTGTTGTTATATACCA 567			
Db	568	AGAGGAGAGATTGTTGTTATATACCA 594			
RESULT 15					
AAA44369					
ID	AAA44369	standard; cDNA; 895 BP.			
XX	AAA44369;				
XX	21-AUG-2000	(first entry)			
DE	Human secreted expressed sequence tag SEQ ID NO:944.				
XX	Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;				
KW	expressed sequence tag; EST; probe; chemotactic; proliferative;				
KW	immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;				
KW	thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;				
KW	antiviral; antidiabetic; antiaesthetic; vulnerary; antiparkinsonian;				
KW	antitumor; osteoprotective; neuroprotective; nociceptive; antipneumatic;				
KW	cerebroprotective; anticonvulsant; antidepressant; gene therapy;				
KW	vaccine; autoimmunity disorder; multiple sclerosis; allergic condition;				
KW	insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;				
KW	lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;				
KW	central nervous system disorder; Alzheimer's disease; stroke;				
KW	Parkinson's disease; Huntington's disease; coagulation disorder;				
KW	haemophilia; thrombosis; inflammatory disorder; Crohn's disease;				
KW	tumour; infection; depression; psoriasis; ss.				
XX	Homo sapiens.				
OS	Wo200021991-A1.				
PN	20-APR-2000.				
XX	15-OCT-1999;	99WO-US24206.			
XX	15-OCT-1998;	98US-0104436.			
XX	(GEM) GENETICS INST INC.				
PI	Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;				
FI	Merberg D, Treacy W, Bowman MR;				
XX	WPI; 2000-317938/27.				
XX	Isolated polynucleotides, and encoded proteins, comprising secreted				
PT	expressed sequence tags (sESTs), useful for treating various disorders				
PT	such as autoimmune, infectious, and central nervous system disorders -				

XX Claim 1; Page 446; 803pp; English.

PS AAA45926 to AAA45925 represent specifically claimed secreted expressed
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
CC tissue sources. The SESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemotactic; proliferative; immunomodulatory; haematopoietic;
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytosratic; antibacterial; antifungal; antiviral; antidiabetic;
CC antiaesthetic; vulnerrary; antitumor; osteopathic; neuroprotective;
CC nociceptive; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antidepressant. The SESTs can be used for gene
CC therapy and in vaccines. The SESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention.

XX
SQ Sequence 895 BP; 274 A; 195 C; 224 G; 201 T; 1 other;

Query Match 96.4%; Score 546.4; DB 21; Length 895;
Best Local Similarity 99.8%; Pred. No. 5.2e-155;

Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 AAGAAACCACTGGAAAAAATGAATTCCTTATCTTGGCAATTTTCGGTGTTC 79
DB 21 AAGAAACCACTGGAAAAAATGAATTCCTTATCTTGGCAATTTTCGGTGTTC 80
QY 80 ACCCTTTATCCCTGTGCTCTGGAAAGCTATATGCAAGATGGCATCTCTAAGAGGACTT 139
DB 81 ACCCTTTATCCCTGTGCTCTGGAAAGCTATATGCAAGATGGCATCTCTAAGAGGACTT 140
QY 140 TTGAAGAAATPAAAGAAATAGCCAGCTGTGAGATGTGCTAAGCAATCATCAACC 199
DB 141 TTGAAGAAATPAAAGAAATAGCCAGCTGTGAGATGTGCTAAGCAATCATCAACC 200
QY 200 TAGCTGTTTATGTAAAGCCCAAGACAGATCTTATGAGCGATTGGCACTTCTGGTTGATA 259
DB 201 TAGCTGTTTATGTAAAGCCCAAGACAGATCTTATGAGCGATTGGCACTTCTGGTTGATA 260
QY 260 CTGTTGGACCCAGACTGAGTGGCTCCAAAGACCTAAGAAAGCCATCCAAATTTATGTACC 319
DB 261 CTGTTGGACCCAGACTGAGTGGCTCCAAAGACCTAAGAAAGCCATCCAAATTTATGTACC 320
QY 320 AAAAAGCTGACAGAGATGGCTGAGAAAGTTCACTTGAGCCAGTGAATACCCCACT 379
DB 321 AAAAAGCTGACAGAGATGGCTGAGAAAGTTCACTTGAGCCAGTGAATACCCCACT 380
QY 380 GGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGAGCCAAAGATTCAATAGATACCATCC 439
DB 381 GGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGAGCCAAAGATTCAATAGATACCATCC 440
QY 440 TGGGCTTTGGAGAGAGCATTTGGGACTCCTCAAGAGGCAATTACAGCAAGTTCTGTGGG 499
DB 441 TGGGCTTTGGAGAGAGCATTTGGGACTCCTCAAGAGGCAATTACAGCAAGTTCTGTGGG 500
QY 500 TGACCTCTTTGATGAATCTGAGAGAAAGGCTCAGAAAGAGGGAAGATTGTTGTTT 559
DB 501 TGACCTCTTTGATGAATCTGAGAGAAAGGCTCAGAAAGAGGGAAGATTGTTGTTT 560
QY 560 ATAACCAA 567
DB 561 ATAACCAA 568

Search completed: December 22, 2003, 15:10:10
Job time : 140.658 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:22:34 ; Search time 1109.8 Seconds
(without alignments)
12417.225 Million cell updates/sec

Title: US-09-745-763-35_COPY_57_623

Perfect score: 567
Sequence: 1 CCGATCAGATTATCTTACA.....AGATTGTTGTTATATACCA 567

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: gb_gss1:*
28: gb_gss2:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
1	567	100.0	873 14	CB984963 AGENCOURT
2	567	100.0	1038 13	BX355940 AGENCOURT
3	567	100.0	1050 13	BX439467 AGENCOURT
4	567	100.0	1126 13	BX360507 AGENCOURT

5	566.2	99.9	919	13	BX355995	BX355995
6	565.4	99.7	1012	13	BX416896	BX416896
7	563.8	99.4	879	13	BQ218838	BQ218838
8	560	98.8	1024	11	BC012019	BC012019
9	556	98.1	806	12	BI764052	BI764052
10	556	98.1	1081	13	BQ072892	BQ072892
11	555	97.9	745	12	BI462584	BI462584
12	554.4	97.8	783	12	BI754468	BI754468
13	554.4	97.8	909	13	BQ878966	BQ878966
14	553.4	97.6	822	12	BI911772	BI911772
15	553.4	97.6	922	12	BG761741	BG761741
16	553.2	97.6	729	12	BI460469	BI460469
17	553.2	97.6	998	13	BX460463	BX460463
18	552.8	97.5	786	10	BG430966	BG430966
19	552.4	97.4	945	12	BI909780	BI909780
20	551.8	97.3	801	12	BI754143	BI754143
21	547	96.5	836	14	CB958603	CB958603
22	544.8	96.1	1201	13	BX379674	BX379674
23	542.6	95.7	722	12	BI766288	BI766288
24	539.2	95.1	944	12	BI819273	BI819273
25	533	94.0	701	12	BI771109	BI771109
26	528.8	93.3	727	12	BG765433	BG765433
27	528	93.1	756	12	BI757436	BI757436
28	526.6	92.9	862	12	BG772207	BG772207
29	524	92.4	849	10	BG166321	BG166321
30	514	90.7	539	14	CB267661	CB267661
31	514	90.7	562	9	AU279837	AU279837
32	506	89.2	889	10	BE906771	BE906771
33	502.8	88.7	945	13	BQ421059	BQ421059
34	480	84.7	669	12	BI910310	BI910310
35	473.6	83.5	736	12	BI769575	BI769575
36	460	81.1	545	14	CA397585	CA397585
37	456.6	80.5	494	12	BI004612	BI004612
38	456	80.4	751	14	CD58931	CD58931
39	452.6	79.8	717	14	CD000772	CD000772
40	450	79.4	866	13	BU146905	BU146905
41	445	78.5	582	10	BG402129	BG402129
42	430.4	75.9	562	14	CB555139	CB555139
43	430.4	75.9	813	14	CD519131	CD519131
44	417	73.5	498	28	AQ129106	AQ129106
45	416	73.4	653	10	BE218907	BE218907

ALIGNMENTS

RESULT 1
CB984963
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CB984963 873 bp mRNA linear EST 01-MAY-2003
AGENCOURT 13567650 NIH-MGC 184 Homo sapiens cDNA clone
IMAGE:30326745 5', mRNA sequence.
CB984963
CB984963.1 GI:30279487
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/
1 (bases 1 to 873)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMU at:
http://image.llnl.gov
Plate: NDCM134 row: k column: 10
High quality sequence stop: 515.

Db 558 ATTCATAGATAGCATCTGGGCTTGGCAGCAGCATTTGGACTCTCCAGAGGCAATT 617

QY 481 ACAGCAGAGTTCTGTGTGACCTTTTCATGAATGCAGAGAGGGCTCAGAAACA 540

Db 618 ACAGCAGAGTTCTGTGTGACCTTTTCATGAATGCAGAGAGGGCTCAGAAACA 677

QY 541 AGAGGGAAGATTGTTTATTAACCA 567

Db 678 AGAGGGAAGATTGTTTATTAACCA 704

RESULT 3
BX439467 1050 bp mRNA linear EST 15-MAY-2003
LOCUS BX439467 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE010YJ06
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX439467 GI:30777746
VERSION BX439467.1 GI:30777746
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1050)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE010DE03QPI&cluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE010DE03QPI.
Location/Qualifiers
FEATURES
source
1..1050
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE010YJ06"
/tissue="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT 307 a 233 c 271 g 239 t

ORIGIN

Query Match 100.0%; Score 567; DB 13; Length 1050;
Best Local Similarity 100.0%; Pred. No. 4,4e-121;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTATCAGATTATCTTAACAAGAAACCACTGAAAAAATGAATTCCTTATCTTC 60

Db 100 CCTATCAGATTATCTTAACAAGAAACCACTGAAAAAATGAATTCCTTATCTTC 159

QY 61 GCATTTTTCGTGTGACCTTTTACCTGTGCTCTGGGAAAGCTATATGCAAGAT 120

Db 160 GCATTTTTCGTGTGACCTTTTACCTGTGCTCTGGGAAAGCTATATGCAAGAT 219

QY 121 GCATCTCTAGAGAGCTTTTGAAGAATAAAGAAATATGACGCTGTGAGATGT 180

Db 220 GCATCTCTAGAGAGCTTTTGAAGAATAAAGAAATATGACGCTGTGAGATGT 279

QY 181 GCTAAGCAATCATCACTAGCTGTTATGTATTAAGCCAGAAAGATCCTATGAGCA 240

Db 280 GCTAAGCAATCATCACTAGCTGTTATGTATTAAGCCAGAAAGATCCTATGAGCA 339

QY 241 TTGGCACTTGTGTTGATCTGTGACCGAGCTGAGTGGCTCCAGAACTTAAAGAAA 300

Db 340 TTGGCACTTGTGTTGATCTGTGACCGAGCTGAGTGGCTCCAGAACTTAAAGAAA 399

QY 301 GCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCACTGAG 360

Db 400 GCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCACTGAG 459

QY 361 CCAATGGAATTAACCCACTGGAGAGGGGAGAAAGATCACTGTGATCTGAGCCAGA 420

Db 460 CCAATGGAATTAACCCACTGGAGAGGGGAGAAAGATCACTGTGATCTGAGCCAGA 519

QY 421 ATTCATAGATAGCATCTGGGCTTGGAGAGCACTTGGACTCTCCAGAAAGCAATT 480

Db 520 ATTCATAGATAGCATCTGGGCTTGGAGAGCACTTGGACTCTCCAGAAAGCAATT 579

QY 481 ACAGCAGAGTTCTGTGTGACCTTTTCATGAATGCAGAGAGGGCTCAGAAACA 540

Db 580 ACAGCAGAGTTCTGTGTGACCTTTTCATGAATGCAGAGAGGGCTCAGAAACA 639

QY 541 AGAGGGAAGATTGTTTATTAACCA 567

Db 640 AGAGGGAAGATTGTTTATTAACCA 666

RESULT 4
BX360507 1126 bp mRNA linear EST 05-MAY-2003
LOCUS BX360507 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1071YA19 5-PRIME, mRNA sequence.
ACCESSION BX360507
VERSION BX360507.1 GI:30374439
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1126)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1071YA19QPI&cluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1071YA19QPI.
Location/Qualifiers
FEATURES
source
1..1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1071YA19"
/tissue="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 311 a 249 c 298 g 246 t 22 others

ORIGIN

Query Match 100.0%; Score 567; DB 13; Length 1126;
Best Local Similarity 100.0%; Pred. No. 4,4e-121;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

Oy 1 CCTATCAGATTATCTTAAAGAAAACCACTGAGAAAAAATGAATTCCTTATCTTC 60
Db 168 CTTATCAGATTATCTTAAAGAAAACCACTGAGAAAAAATGAATTCCTTATCTTC 227
Oy 61 GCATTTTTCGGTGTGTGTACCTTTTATCCCTGTGCTCTGAGAAAAGCTATATGCAAGAT 120
Db 228 GCATTTTTCGGTGTGTGTACCTTTTATCCCTGTGCTCTGAGAAAAGCTATATGCAAGAT 287
Oy 121 GGCATCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGACCTGTGAGATTTT 180
Db 288 GGCATCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGACCTGTGAGATTTT 347
Oy 181 GCTAAAGCAATCATCAACCTAGCTGTATGTTAAAGCCAGAACAGATCCTATGAGCA 240
Db 348 GCTAAAGCAATCATCAACCTAGCTGTATGTTAAAGCCAGAACAGATCCTATGAGCA 407
Oy 241 TTGGCACTTCTGTGTATCTGTGGACCCAGACTGAGTGGCTCCAGAACCTAGAAAA 300
Db 408 TTGGCACTTCTGTGTATCTGTGGACCCAGACTGAGTGGCTCCAGAACCTAGAAAA 467
Oy 301 GCCATCCAAATTTATGTACCAAAACCTGCAGCAAGATGGGCTGAGAAAAGTTCACTGAG 360
Db 468 GCCATCCAAATTTATGTACCAAAACCTGCAGCAAGATGGGCTGAGAAAAGTTCACTGAG 527
Oy 361 CCAGTGAAGATTAACCCACTGGGAGAGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 420
Db 528 CCAGTGAAGATTAACCCACTGGGAGAGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 587
Oy 421 ATTCTAAGATAGGCAATCCTGGGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAAT 480
Db 588 ATTCTAAGATAGGCAATCCTGGGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAAT 647
Oy 481 ACAGCAAGAAATTTCTGTGTGTACCTCTTCCATGAATCTGCAGAAAGGGCTCAGAAACA 540
Db 648 ACAGCAAGAAATTTCTGTGTGTGTGACCTCTTCCATGAATCTGCAGAAAGGGCTCAGAAACA 707
Oy 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
Db 708 AGAGGGAAGATTGTTGTTTATTAACCA 734

```

```

RESULT 5          919 bp  mRNA  linear  EST 02-MAY-2003
BX335995          Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
LOCUS             BX335995
DEFINITION        BX335995 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
ACCESSION         BX335995
VERSION           BX335995.1 GI:30339459
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE         1 (bases 1 to 919)
AUTHORS           Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE             Full-length cDNA libraries and normalization
JOURNAL           Unpublished
COMMENT           Contact: Genoscope
                  Genoscope - Centre National de Sequencage
                  BP 191 91006 Evry cedex - France
                  Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
                  Library was constructed by Life Technologies, a division of
                  Invitrogen. This sequence belongs to sequence cluster 4663.f For
                  more information about this cluster, see
                  http://www.genoscope.cns.fr/
                  cgl-bi/cluster.cgi?seq=CS0D1022DE080P1c&cluster=4663.f. Contact :
                  Feng Liang Email : fliang@lifetech.com URL :
                  http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
                  Faraday Avenue Genoscope sequence ID : CS0D1022DE080P1.
                  Location/Qualifiers
                  1..919
                  /organism="Homo sapiens"

```

FEATURES

source

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1022YJ16"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      273 a      206 c      226 g      208 t      6 others
ORIGIN

```

```

Query Match      99.9%; Score 566.2; DB 13; Length 919;
Best Local Similarity 99.6%; Pred. No. 6.8e-121;
Matches 565; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 CCTATCAGATTATCTTAAAGAAAACCACTGAGAAAAAATGAATTCCTTATCTTC 60
Db 84 CTTATCAGATTATCTTAAAGAAAACCACTGAGAAAAAATGAATTCCTTATCTTC 143
Oy 61 GCATTTTTCGGTGTGTGTACCTTTTATCCCTGTGCTCTGAGAAAAGCTATATGCAAGAT 120
Db 144 GCATTTTTCGGTGTGTGTGTACCTTTTATCCCTGTGCTCTGAGAAAAGCTATATGCAAGAT 203
Oy 121 GGCATCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGACCTGTGAGATTTT 180
Db 204 GGCATCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGACCTGTGAGATTTT 263
Oy 181 GCTAAAGCAATCATCAACCTAGCTGTATGTTAAAGCCAGAACAGATCCTATGAGCA 240
Db 264 GCTAAAGCAATCATCAACCTAGCTGTATGTTAAAGCCAGAACAGATCCTATGAGCA 323
Oy 241 TTGGCACTTCTGTGTATCTGTGGACCCAGACTGAGTGGCTCCAGAACCTAGAAAA 300
Db 324 TTGGCACTTCTGTGTATCTGTGGACCCAGACTGAGTGGCTCCAGAACCTAGAAAA 383
Oy 301 GCCATCCAAATTTATGTACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAG 360
Db 384 GCCATCCAAATTTATGTACCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGAG 443
Oy 361 CCAGTGAAGATTAACCCACTGGGAGAGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 420
Db 444 CCAGTGAAGATTAACCCACTGGGAGAGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 503
Oy 421 ATTCTAAGATAGGCAATCCTGGGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAAT 480
Db 504 ATTCTAAGATAGGCAATCCTGGGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAAT 563
Oy 481 ACAGCAAGAAATTTCTGTGTGTACCTCTTCCATGAATCTGCAGAAAGGGCTCAGAAACA 540
Db 564 ACAGCAAGAAATTTCTGTGTGTGTGACCTCTTCCATGAATCTGCAGAAAGGGCTCAGAAACA 623
Oy 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
Db 624 AGAGGGAAGATTGTTGTTTATTAACCA 650

```

```

RESULT 6          1012 bp  mRNA  linear  EST 13-MAY-2003
BX416896          Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE001YF04
LOCUS             BX416896
DEFINITION        BX416896 Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE001YF04
ACCESSION         BX416896
VERSION           BX416896.1 GI:30650313
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE         1 (bases 1 to 1012)
AUTHORS           Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE             Full-length cDNA libraries and normalization
JOURNAL           Unpublished

```

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE001D02Q2P1c&cluster=4663.f. Contact :
Peng Liang Email : liliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0DE001D02Q2P1.

FEATURES

source

1. 1012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE001F04"
/issue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
Library was not normalized."
BASE COUNT 281 a 240 c 275 g 215 t 1 others
ORIGIN

Query Match 99.7%; Score 565.4; DB 13; Length 1012;
Best Local Similarity 99.8%; Pred. No. 1e-120;
Matches 566; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 60
213 CTTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 272
61 GCATTTTCGGTGGTGTCACTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
273 GCATTTTCAGTGGTGTCACTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 332
121 GGCATCTCTAAGAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT 180
333 GGCATCTCTAAGAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT 392
181 GCTAAGCAATCATCAACTGAGCTGTTATGTTAAGCCGAGACAGATCTTATGAGCGA 240
393 GCTAAGCAATCATCAACTGAGCTGTTATGTTAAGCCGAGACAGATCTTATGAGCGA 452
241 TTGGCACTTCTGTTGATTAAGTGTGAGCCAGACTGAGTGGCTCCAGAAACCTAGAAAA 300
453 TTGGCACTTCTGTTGATTAAGTGTGAGCCAGACTGAGTGGCTCCAGAAACCTAGAAAA 512
301 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAAATTCACCTGAG 360
513 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAAATTCACCTGAG 572
361 CCAAGTGAATATACCCCACTGGAGAGGGGAGAGAAATAGCTGTGATCTGAGCCAGAA 420
573 CCAAGTGAATATACCCCACTGGAGAGGGGAGAGAAATAGCTGTGATCTGAGCCAGAA 632
421 ATTGATATAGATAGCATCTGAGGCTTGGAGAGAGATTTGGGACCTCTCCAGAAAGGCA 480
633 ATTGATATAGATAGCATCTGAGGCTTGGAGAGAGATTTGGGACCTCTCCAGAAAGGCA 692
481 ACAGCAAGAGTTCTGTGTGAGCTCTTTGATGAACTGCAAGAAAGGCTCAGAAACA 540
693 ACAGCAAGAGTTCTGTGTGAGCTCTTTGATGAACTGCAAGAAAGGCTCAGAAACA 752
541 AGAGGAGAAATTTGTTTATTAACCA 567
753 AGAGGAGAAATTTGTTTATTAACCA 779

RESULT 7

LOCUS

DEFINITION

B0218838 879 bp mRNA linear EST 02-MAY-2002
AGENCOURT 7546057 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6025034

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1. 879
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6025034"
/issue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 239 a 216 c 235 g 189 t
ORIGIN

FEATURES

source

1 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 60
103 CTTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 162
61 GCATTTTCGGTGGTGTCACTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
163 GCATTTTCGGTGGTGTCACTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 222
121 GGCATCTCTAAGAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT 180
223 GGCATCTCTAAGAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT 282
181 GCTAAGCAATCATCAACTGAGCTGTTATGTTAAGCCGAGACAGATCTTATGAGCGA 240
283 GCTAAGCAATCATCAACTGAGCTGTTATGTTAAGCCGAGACAGATCTTATGAGCGA 342
241 TTGGCACTTCTGTTGATTAAGTGTGAGCCAGACTGAGTGGCTCCAGAAACCTAGAAAA 300
343 TTGGCACTTCTGTTGATTAAGTGTGAGCCAGACTGAGTGGCTCCAGAAACCTAGAAAA 402
301 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAAATTCACCTGAG 360
403 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAAATTCACCTGAG 462
361 CCAAGTGAATATACCCCACTGGAGAGGGGAGAGAAATAGCTGTGATCTGAGCCAGAA 420

Db 463 CCAGTGAAGATACCCCACTGGAGAGGGAGAGAGATCAGCTGTGATCTGGAGCCAAAGA 522
Oy 421 ATTCAATAGATAGCCATCTGGGCTCTTGAGACAGCATTTGGGAGCTCTCCGAAAGGCATT 480
Db 523 ATTCTATAGATAGCCATCTGGGCTCTTGAGACAGCATTTGGGAGCTCTCCGAAAGGCATT 582
Oy 481 ACAGCAGAAAGTTCTGGTGGTGAACCTTTTCGATGTAATGCAAGAGAGGCGCTCAGAAACA 540
Db 583 ACAGCAGAAAGTTCTGGTGGTGAACCTTTTCGATGTAATGCAAGAGAGGCGCTCAGAAATC 642
Oy 541 AGAGGGAAGATTGTTGTTTATACCA 567
Db 643 AGAGGGAAGATTGTTGTTTATACCA 669

RESULT 8
LOCUS BC012019 1024 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, similar to plasma glutamate carboxypeptidase, clone
IMAGE:4455631, mRNA.
BC012019
ACCESSION BC012019.1 GI:15080558
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1024)
Strausberg, R.
Direct Submission
Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.mci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarane, P.H., Garcia, A.M., Lu, X., Huiy, S.W., Louisseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nantavali,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 28 Row: b Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5174626
This clone has the following problem: retained intron.
Location/Qualifiers
1..1024
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4455631"
/tissue_type="Kidney, hypernephroma"
/clone_id="NIH_MGC_69"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 317 a 210 c 256 g 241 t
ORIGIN

Query Match 98.8%; Score 560; DB 11; Length 1024;
Best local Similarity 100.0%; Pred. No. 1.9e-119;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 GATTATCTTAACAAGAAACCACTGGAAGAAAAAATGAAATTCCTATCTTCGATTTT 67

Db 1 GATTATCTTAACAAGAAACCACTGGAAGAAAAAATGAAATTCCTATCTTCGATTTT 60
Oy 68 TCGGTGGTTCACCTTTTATCCCTGTCCTGGGAAGCATATGCAAGATGGCATCT 127
Db 61 TCGGTGGTTCACCTTTTATCCCTGTCCTGGGAAGCATATGCAAGATGGCATCT 120
Oy 128 CTAAGAGGACTTTGGAAGAAATAAAGAAATATAGCAGCTGGAGATGTTCTAAAG 187
Db 121 CTAAGAGGACTTTGGAAGAAATAAAGAAATATAGCAGCTGGAGATGTTCTAAAG 180
Oy 188 CAATCATCAACCTAGCTGTTTATGTAAGCCGAAACAGATCTATGAGCATGGCAC 247
Db 181 CAATCATCAACCTAGCTGTTTATGTAAGCCGAAACAGATCTATGAGCATGGCAC 240
Oy 248 TTCTGGTGTATACCTGTGGACCCAGACTGAGTGGCTCCAGAACTTGAAGAAACCTCC 307
Db 241 TTCTGGTGTATACCTGTGGACCCAGACTGAGTGGCTCCAGAACTTGAAGAAACCTCC 300
Oy 308 AAATTATGTACCAAAACCTGAGCAAGATGAGGCTGAGAAAGTTCACTGGAGCGATGA 367
Db 301 AAATTATGTACCAAAACCTGAGCAAGATGAGGCTGAGAAAGTTCACTGGAGCGATGA 360
Oy 368 GAATATCCCACTGGGAGAGGGAGAGAAATCAGCTGTGATGCTGAGCCAAAGATTGATA 427
Db 361 GAATATCCCACTGGGAGAGGGAGAGAAATCAGCTGTGATGCTGAGCCAAAGATTGATA 420
Oy 428 AGATAGCCATCTCTGGGCTCTTGAGCAGCAGCATTTGGGATCTCTCCAGAAAGCATTA 487
Db 421 AGATAGCCATCTCTGGGCTCTTGAGCAGCAGCATTTGGGATCTCTCCAGAAAGCATTA 480
Oy 488 AAGTTCTGAGTGAACCTCTTCATGAACTGCAAGAGAGGCGCTCAGAAAGCAAGAGGA 547
Db 481 AAGTTCTGAGTGAACCTCTTCATGAACTGCAAGAGAGGCGCTCAGAAAGCAAGAGGA 540
Oy 548 AGATTGTTGTTTATACCA 567
Db 541 AGATTGTTGTTTATACCA 560

RESULT 9
LOCUS B1764052 806 bp mRNA linear EST 25-SEP-2001
DEFINITION 603043258F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183559 5',
mRNA sequence.
B1764052
ACCESSION B1764052.1 GI:15755630
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 806)
NIH-MGC <http://mgc.mci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11458 row: c column: 16
High quality sequence stop: 804.
Location/Qualifiers
1..806
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5183559"

/lab_host="DH10B"
/clone_1ib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon, 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC Library."

BASE COUNT
ORIGIN

221 a 194 c 213 g 178 t

Query Match 98.1%; Score 556; DB 12; Length 806;
Best Local Similarity 99.8%; Pred. No. 1.6e-118;
Matches 567; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CCTATCAGATTATCTTAACAGAAACCACTGGAAAAAATGAATTCTTATCTTC 60
DB 115 CCTATCAGATTATCTTAACAGAAACCACTGGAAAAAATGAATTCTTATCTTC 174
QY 61 GCATTTTTCGGTGTGTTCACTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAT 120
DB 175 GCATTTTTCGGTGTGTTCACTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAT 234
QY 121 GGCATCTCTAAGAGACTTTTGAAGAAATTAAGAAATATGCACTGTGAGATGTT 180
DB 235 GGCATCTCTAAGAGACTTTTGAAGAAATTAAGAAATATGCACTGTGAGATGTT 294
QY 181 GCTAAGGCAATCATCAACCTAGCTGTATGTAAAGCCAGAAACAGATCTTATGAGCGA 240
DB 295 GCTAAGGCAATCATCAACCTAGCTGTATGTAAAGCCAGAAACAGATCTTATGAGCGA 354
QY 241 TTGGCACTTCTGTTGATCTGTGACCACTGATGAGCTTCAAGAACTAGAAATA 300
DB 355 TTGGCACTTCTGTTGATCTGTGACCACTGATGAGCTTCAAGAACTAGAAATA 414
QY 301 GCCATCCAAATTATGTAACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTGGAG 360
DB 415 GCCATCCAAATTATGTAACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTGGAG 474
QY 361 CCAGTAGATATACCCCACTGGAGAGAGGAGAAAGAAATCAGCTGTGATGCTGAGACCAAGA 420
DB 475 CCAGTAGATATACCCCACTGGAGAGAGGAGAAAGAAATCAGCTGTGATGCTGAGACCAAGA 534
QY 421 ATTCAATAGATAGCATCTGGGTCTTGGAGAGCAGCAATTGGGAACTCTCCAGAAAGGCAATT 480
DB 535 ATTCAATAGATAGCATCTGGGTCTTGGAGAGCAGCAATTGGGAACTCTCCAGAAAGGCAATT 594
QY 481 ACAGCAAGAGTTCTGTGTGACCTCTTTCGATGAAGTGC-AGAGAAAGGCTCAGAAAGC 539
DB 595 ACAGCAAGAGTTCTGTGTGACCTCTTTCGATGAAGTGCAGAGAGAAAGGCTCAGAAAGC 654
QY 540 AAGAGGGAAGATGTTGTTTATTAACCA 567
DB 655 AAGAGGGAAGATGTTGTTTATTAACCA 682

RESULT 10
LOCUS B0072892 1081 bp mRNA linear EST 02-APR-2002
DEFINITION ABEHCOURT_6763138 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5756387
ACCESSION B0072892
VERSION B0072892.1 GI:19901938
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1081)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM412796 row: 0 column: 12
High quality sequence stop: 658.

FEATURES

Source

Location/Qualifiers

1..1081

/organism="Homo sapiens"

/db_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5756387"

/issue_type="leukocyte"

/lab_host="DH10B"

/clone_1ib="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC Library."

BASE COUNT 319 a 256 c 275 g 225 t 2 others

ORIGIN

Query Match 98.1%; Score 556; DB 13; Length 1081;
Best Local Similarity 99.8%; Pred. No. 1.6e-118;
Matches 567; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CCTATCAGATTATCTTAACAGAAACCAACT-GGAAAAAATGAATTCTTATCTT 59
DB 123 CCTATCAGATTATCTTAACAGAAACCAACTGGAAAAAATGAATTCTTATCTT 182
QY 60 CGATTTTTCGGTGTGTTCACTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAA 119
DB 183 CGATTTTTCGGTGTGTTCACTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAA 242
QY 120 TGGCATCTCTAAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT 179
DB 243 TGGCATCTCTAAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT 302
QY 180 TGGTAAAGCAATCATCAACCTAGCTGTATGTGTAAGCCAGAAACAGATCTATGAGCG 239
DB 303 TGGTAAAGCAATCATCAACCTAGCTGTATGTGTAAGCCAGAAACAGATCTATGAGCG 362
QY 240 ATTGGCACTTCTGTGTATCTGTGGAACCAACTGAGTGGCTCCAGAAACTTAAGAAA 299
DB 363 ATTGGCACTTCTGTGTATCTGTGGAACCAACTGAGTGGCTCCAGAAACTTAAGAAA 422
QY 300 AGCCATCCAAATTATGTAACCAAAACCTGACAGAAATGGGCTGGAGAAAGTTCACTGGA 359
DB 423 AGCCATCCAAATTATGTAACCAAAACCTGACAGAAATGGGCTGGAGAAAGTTCACTGGA 482
QY 360 GCCAGTAGAATACCCCACTGGAGAGAGGAGAGAAATGAGCTGTGATGCTGAGAGCAAG 419
DB 483 GCCAGTAGAATACCCCACTGGAGAGAGGAGAGAAATGAGCTGTGATGCTGAGAGCAAG 542
QY 420 AATTCAATAGATAGCATCTGGGTCTTGGAGACAGATTTGGAGACTCTCCAGAAAGGCAAT 479
DB 543 AATTCAATAGATAGCATCTGGGTCTTGGAGACAGATTTGGAGACTCTCCAGAAAGGCAAT 602
QY 480 TACAGCAAGAGTTCTGTGTGACCTCTTTCGATGAAGCTCAGAGAAAGGCTCAGAAAGC 539

```
Db 603 TACGACGAAGTCTGTGGTGGACCTCTTTCGATGAACTGACAGAGAGGCGCTAGAGC 662
Qy 540 AAGAGGGAAGATTGTTTATTAACCA 567
Db 663 AAGAGGGAAGATTGTTTATTAACCA 690

RESULT 11
LOCUS B1462584 745 bp mRNA linear EST 21-AUG-2001
DEFINITION 60302005F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5268937 5',
mRNA sequence.
ACCESSION B1462584
VERSION B1462584.1 GI:15253240
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 745)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Mikhail Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M1678 row: i column: 02
High quality sequence stop: 741.
Location/Qualifiers
1. 745
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5268937"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (Gcgag
); Oligo-dT primed using primer 5'-TTTATTTTATTTTATTTT-3',
size-selected for average insert size 2.2 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 213 a 169 c 198 g 164 t 1 others
ORIGIN
Query Match 97.9%; Score 555; DB 12; Length 745;
Best Local Similarity 99.8%; Pred. No. 2.7e-118;
Matches 565; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CCGATCAGATTATCTTAACAAGAAACCAACTGAAAAAAGAAATCCTATATCTC 60
Db 87 CCGATCAGATTATCTTAACAAGAAACCAACTGAAAAAAGAAATCCTATATCTC 146
Qy 61 GCATTTTCGGTGTGTCACTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAT 120
Db 147 GCATTTTCGGTGTGTCACTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAT 206
Qy 121 GGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGAGATGTT 180
Db 207 GGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGAGATGTT 266
Qy 181 GCTAAGCAATCATCACTAGCTGTTTATAGTAAAGCCCAAGACATCCTATAGCGA 240
```

```
Db 267 GCTAAGCAATCATCACTAGCTGTTTATAGTAAAGCCCAAGACATCCTATAGCGA 326
Qy 241 TTGGCACTCTGTGTTATATCTGTGGACCCAGACTAGTGGCTCCAAAGAACTAGAAAA 300
Db 327 TTGGCACTCTGTGTTATATCTGTGGACCCAGACTAGTGGCTCCAAAGAACTAGAAAA 386
Qy 301 GCCATCAATATATATGTAACCAAACTGCAAGCAAGATGGGCTGGAGAAAGTTCACTGGAG 360
Db 387 GCCATCAATATATATGTAACCAAACTGCAAGCAAGATGGGCTGGAGAAAGTTCACTGGAG 446
Qy 361 CCAAGTAAGATATCCCACTGGAGAGAGGAGAAAGAAATCACTGTGATGCTGAGCCCAAG 420
Db 447 CCAAGTAAGATATCCCACTGGAGAGAGGAGAAAGAAATCACTGTGATGCTGAGCCCAAG 506
Qy 421 ATTCATTAAGATAGCCATCTCTGGTCTTTGGCAGAGCATTTGGACCTCTCCAGAAAGCAT 480
Db 507 ATTCATTAAGATAGCCATCTCTGGTCTTTGGCAGAGCATTTGGACCTCTCCAGAAAGCAT 566
Qy 481 ACAGCAAGATGTTGTTGGTGAACCTCTTTCGATGAATGAGAGAGAGGCTCAGAGCA 540
Db 567 ACAGCAAGATGTTGTTGGTGAACCTCTTTCGATGAATGAGAGAGAGGCTCAGAGCA 626
Qy 541 AGAGGGAAGATTGTTTATTAACCA 567
Db 627 AGAGGGAAGATTG-TGTTTATTAACCA 652

RESULT 12
LOCUS B1754468 783 bp mRNA linear EST 25-SEP-2001
DEFINITION 603022607F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5193385 5',
mRNA sequence.
ACCESSION B1754468
VERSION B1754468.1 GI:15746046
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 783)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M1483 row: m column: 02
High quality sequence stop: 783.
Location/Qualifiers
1. 783
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5193385"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
BASE COUNT 223 a 188 c 197 g 175 t
```

ORIGIN

Query Match 97.8%; Score 554.4; DB 12; Length 793;
 Best Local Similarity 99.6%; Pred. No. 3.7e-118;
 Matches 566; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CCTATCAGATTATCTTAAACAAGAAAAACCACTGAGAAAAAATGAAATTCCTTATCTTC 60
 DB 73 CCTATCAGATTATCTTAAACAAGAAAAACCACTGAGAAAAAATGAAATTCCTTATCTTC 132
 QY 61 GCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGAGAAAGCTATATGCAAGAT 120
 DB 133 GCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGAGAAAGCTATATGCAAGAT 192
 QY 121 GGCATCTCTAAGAGGACTTTTGAAGAATAAAGAAATAGCAGCTGTGAGATGT 180
 DB 193 GGCATCTCTAAGAGGACTTTTGAAGAATAAAGAAATAGCAGCTGTGAGATGT 252
 QY 181 GCTAAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCCTATGAGCA 240
 DB 253 GCTAAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCCTATGAGCA 312
 QY 241 TTGGCATCTTGGTGTGATCTGTGAGACCACTGATGCTGCCAAGACTTGAAGAAA 300
 DB 313 TTGGCATCTTGGTGTGATCTGTGAGACCACTGATGCTGCCAAGACTTGAAGAAA 372
 QY 301 GCCATCCAAATTATGTATCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGAG 360
 DB 373 GCCATCCAAATTATGTATCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGAG 432
 QY 361 CCAGTGAAGATATACCCCACTGGAGAGAGGAGAGAAATCAGCTGTATGCTGAGCCAGA 420
 DB 433 CCAGTGAAGATATACCCCACTGGAGAGAGGAGAGAAATCAGCTGTATGCTGAGCCAGA 492
 QY 421 ATTCAATGAATAGCCATCTTGGGCTTTGGCAGCAGCATTTGGAGCTCTCCAGAAAGCAT 480
 DB 493 ATTCAATGAATAGCCATCTTGGGCTTTGGCAGCAGCATTTGGAGCTCTCCAGAAAGCAT 552
 QY 481 ACAGAGAAAGTTGGTGGTGTGACCTTTGATGTAAGTGCAGAGAAAGGCTCCAGAAACA 540
 DB 553 ACAGAGAAAGTTGGTGGTGTGACCTTTGATGTAAGTGCAGAGAAAGGCTCCAGAAACA 612
 QY 541 AGAGGGAAG-ATTGTTGTTTATTAACCA 567
 DB 613 AGAGGGAAGCATTTGCTTTATTAACCA 640

RESULT 13
 B0878966 909 bp mRNA linear EST 16-AUG-2002
 LOCUS B0878966
 DEFINITION AGENCOURT 8183107 lupski_dorsal_root_ganglion Homo sapiens cDNA
 clone IMAGE:6184129 5', mRNA sequence.
 ACCESSION B0878966
 VERSION B0878966.1 GI:22270974
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgarbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.lnl.gov
 Plate: LLM13572 row: n column: 02

FEATURES
 source
 High quality sequence stop: 696.
 Location/Qualifiers
 1..909
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6184129"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /clone_id="lupski_dorsal_root_ganglion"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
 Note: Site 2: SalI; cDNA made by oligo-dT priming.
 directionally cloned using the following adaptors:
 5'-TCGACCCAGGCGCTCCG-3' and
 5'-GACTAGTTCTAGATGCGAGCGGCCCTT(15)-3'. Size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

BASE COUNT 273 a 202 c 226 g 207 t 1 others

ORIGIN

Query Match 97.8%; Score 554.4; DB 13; Length 909;
 Best Local Similarity 99.6%; Pred. No. 3.7e-118;
 Matches 566; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CCTATCAGATTATCTTAAACAAGAAACCAACT-GGAAAAAATGAAATTCCTTATCTT 59
 DB 21 CCTATCAGATTATCTTAAACAAGAAACCAACTGGAAAAAATGAAATTCCTTATCTT 80
 QY 60 CGCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGAGAAAGCTATATGCAAGAA 119
 DB 81 CGCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGAGAAAGCTATATGCAAGAA 140
 QY 120 TTGGCATCTTGAAGAGGACTTTTGAAGAAATAAAGAAATATACCGCTGTGAGATGT 179
 DB 141 TTGGCATCTTGAAGAGGACTTTTGAAGAAATAAAGAAATATACCGCTGTGAGATGT 200
 QY 141 TTGGCATCTTGAAGAGGACTTTTGAAGAAATAAAGAAATATACCGCTGTGAGATGT 200
 QY 180 TGCTAAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCTATGAGCG 239
 DB 201 TGCTAAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCTATGAGCG 260
 QY 240 ATTGGCACTTCTGTTGATCTGTTGACCCAGACTGATGCTCCAGAAACTTGAAGAA 299
 DB 261 ATTGGCACTTCTGTTGATCTGTTGACCCAGACTGATGCTCCAGAAACTTGAAGAA 320
 QY 300 AGCCATCCAAATTATGTATACCAAAACCTGACAGCAAGATGGGTGAGAAAGTTCACTGGA 359
 DB 321 AGCCATCCAAATTATGTATACCAAAACCTGACAGCAAGATGGGTGAGAAAGTTCACTGGA 380
 QY 360 GCCAGTGAATATACCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGAGCCAG 419
 DB 381 GCCAGTGAATATACCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGAGCCAG 440
 QY 420 AATTCATTAAGTATAGCCATCTGAGTCTTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAT 479
 DB 441 AATTCATTAAGTATAGCCATCTGAGTCTTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAT 500
 QY 480 TACAGCAAGAACTTCTGGTGTGACCTCTTTCGATGAAGAACTGAGAAAGGCTCAGAAAGC 539
 DB 501 TACAGCAAGAACTTCTGGTGTGACCTCTTTCGATGAAGAACTGAGAAAGGCTCAGAAAGC 560
 QY 540 AAGAGGGAAGATTGTTGTTTATTAACCA 567
 DB 561 AAGAGGGAAGATTGTTGTTTATTAACCA 588

RESULT 14
 B1911772 822 bp mRNA linear EST 16-OCT-2001
 LOCUS B1911772

DEFINITION 603065377F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214451 5', mRNA sequence.

ACCESSION B1911772

VERSION B1911772.1 GI:16175644

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgi.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHAM1538 row: J column: 20
High quality sequence start: 31
High quality sequence stop: 795.
Location/Qualifiers

FEATURES

source

1..822

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5214451"

/tissue_type="leukocyte"

/lab_host="DH10B"

/clone_id="NIH_MGC_118"

/note="Vector: pCMV-SPORT6, Site 1: NotI, Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC Library."

BASE COUNT 232 a 192 c 215 g 183 t

ORIGIN

Query Match 97.6%; Score 553.4; DB 12; Length 922;
Best Local Similarity 99.6%; Pred. No. 6.4e-118;
Matches 565; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

1 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60
2 |||||||
3 |||||||
4 |||||||
5 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 154
6 |||||||
7 |||||||
8 |||||||
9 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
10 |||||||
11 |||||||
12 |||||||
13 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 214
14 |||||||
15 |||||||
16 |||||||
17 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 214
18 |||||||
19 |||||||
20 |||||||
21 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 180
22 |||||||
23 |||||||
24 |||||||
25 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 274
26 |||||||
27 |||||||
28 |||||||
29 GCTAAGACATCATCACTAGCTGTGTTATGTTAAAGCCAGACAGATCCATAGAGCGA 240
30 |||||||
31 |||||||
32 |||||||
33 GCTAAGACATCATCACTAGCTGTGTTATGTTAAAGCCAGACAGATCCATAGAGCGA 334
34 |||||||
35 |||||||
36 |||||||
37 TTGGCACTTCTGTTGATCTGTGGACCCAGACTGATGCTCCAGAACCTAGAAAAA 300
38 |||||||
39 |||||||
40 |||||||
41 TTGGCACTTCTGTTGATCTGTGGACCCAGACTGATGCTCCAGAACCTAGAAAAA 394
42 |||||||
43 |||||||
44 |||||||
45 GCATTCCTTAAGTATGATCAAAACCTGACGAGAGATGGCGCTGAGAGAAATTCCTGGAG 360
46 |||||||
47 |||||||
48 |||||||
49 GCATTCCTTAAGTATGATCAAAACCTGACGAGAGATGGCGCTGAGAGAAATTCCTGGAG 454
50 |||||||

361 CCACTGAGAAATACCCCACTGGAGAGAGGAGAAATCACTGATGATGAGCCAGA 420
362 |||||||
363 |||||||
364 |||||||
365 CCACTGAGAAATACCCCACTGGAGAGAGGAGAAATCACTGATGATGAGCCAGA 514
366 |||||||
367 |||||||
368 |||||||
369 ATTCAATGAATAGGCACTCTGGGCTTTGGAGAGCATTTGGAGCTCTCCAGAGGCA 480
370 |||||||
371 |||||||
372 |||||||
373 ATTCAATGAATAGGCACTCTGGGCTTTGGAGAGCATTTGGAGCTCTCCAGAGGCA 574
374 |||||||
375 |||||||
376 |||||||
377 ACACAGAAATGTTCTGGTGTGATCCTCTTTGATGATGATGAGAGAGAGGCTCAGAGCA 540
378 |||||||
379 |||||||
380 |||||||
381 ACACAGAAATGTTCTGGTGTGATCCTCTTTGATGATGATGAGAGAGAGGCTCAGAGCA 634
382 |||||||
383 |||||||
384 |||||||
385 ACACAGAAATGTTCTGGTGTGATCCTCTTTGATGATGATGAGAGAGAGGCTCAGAGCA 634
386 |||||||
387 |||||||
388 |||||||
389 AGAGGAAAGATTTGTTTATTAACCA 567
390 |||||||
391 |||||||
392 AGAGGAAAGATTCG-TGTTTATTAACCA 660
393 |||||||

RESULT 15
BG761741
LOCUS

DEFINITION BG761741 922 bp mRNA linear EST 15-MAY-2001
602717936F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4841505 5', mRNA sequence.

ACCESSION BG761741

VERSION BG761741.1 GI:14072394

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgi.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM1674 row: G column: 10
High quality sequence stop: 796.
Location/Qualifiers

1..922

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4841505"

/tissue_type="melanotic melanoma, high MDR (cell line)"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH_MGC_49"

/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRV/XhoI sites using the following 5' adaptor: GCGACAGAG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 279 a 204 c 227 g 211 t 1 others

ORIGIN

Query Match 97.6%; Score 553.4; DB 12; Length 922;
Best Local Similarity 99.6%; Pred. No. 6.4e-118;
Matches 565; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

1 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60
2 |||||||
3 |||||||
4 |||||||
5 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 154
6 |||||||
7 |||||||
8 |||||||
9 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
10 |||||||
11 |||||||
12 |||||||
13 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 214
14 |||||||
15 |||||||
16 |||||||
17 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 214
18 |||||||
19 |||||||
20 |||||||
21 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 180
22 |||||||
23 |||||||
24 |||||||
25 GCATTTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 274
26 |||||||
27 |||||||
28 |||||||
29 GCTAAGACATCATCACTAGCTGTGTTATGTTAAAGCCAGACAGATCCATAGAGCGA 240
30 |||||||
31 |||||||
32 |||||||
33 GCTAAGACATCATCACTAGCTGTGTTATGTTAAAGCCAGACAGATCCATAGAGCGA 334
34 |||||||
35 |||||||
36 |||||||
37 TTGGCACTTCTGTTGATCTGTGGACCCAGACTGATGCTCCAGAACCTAGAAAAA 300
38 |||||||
39 |||||||
40 |||||||
41 TTGGCACTTCTGTTGATCTGTGGACCCAGACTGATGCTCCAGAACCTAGAAAAA 394
42 |||||||
43 |||||||
44 |||||||
45 GCATTCCTTAAGTATGATCAAAACCTGACGAGAGATGGCGCTGAGAGAAATTCCTGGAG 360
46 |||||||
47 |||||||
48 |||||||
49 GCATTCCTTAAGTATGATCAAAACCTGACGAGAGATGGCGCTGAGAGAAATTCCTGGAG 454
50 |||||||


```
QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAGAT 120
    |||||
Db 82 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAGAT 141
    |||||
QY 121 GGCATCTCTTAAGAGGACTTTTGAAGAAATTAAGAGAAATAGCCAGCTGTGAGATGT 180
    |||||
Db 142 GGCATCTCTTAAGAGGACTTTTGAAGAAATTAAGAGAAATAGCCAGCTGTGAGATGT 201
    |||||
QY 181 GCTAAGCAATCAATCAACCTAGCTGTTTATGTATAGCCAGAAACAGATCCTATAGGGA 240
    |||||
Db 202 GCTAAGCAATCAATCAACCTAGCTGTTTATGTATAGCCAGAAACAGATCCTATAGGGA 261
    |||||
QY 241 TTGGCACTTCTGTGTGATCTGTGACCCAGACTGAGTGGCTCCAGAACTTAGAAAA 300
    |||||
Db 262 TTGGCACTTCTGTGTGATCTGTGACCCAGACTGAGTGGCTCCAGAACTTAGAAAA 321
    |||||
QY 301 GCCATCCAAATTATGTACCAAAACCTGACAGCAAGATGGCTGAGAAAATTACCTGGAG 360
    |||||
Db 322 GCCATCCAAATTATGTACCAAAACCTGACAGCAAGATGGCTGAGAAAATTACCTGGAG 381
    |||||
QY 361 CCAAGTGAATAACCCCACTGGGAGAGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 420
    |||||
Db 382 CCAAGTGAATAACCCCACTGGGAGAGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 441
    |||||
QY 421 ATTCATAGATAGCCATCTGGGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAT 480
    |||||
Db 442 ATTCATAGATAGCCATCTGGGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAT 501
    |||||
QY 481 ACAGCAGAAGTCTGTGTGTGACCTTTTCATGAACTGCAGAGAAAGGCTCAGAAACA 540
    |||||
Db 502 ACAGCAGAAGTCTGTGTGTGACCTTTTCATGAACTGCAGAGAAAGGCTCAGAAACA 560
    |||||
QY 541 AGAGGAAAGATTGTGTTTATTAACCA 567
    |||||
Db 561 AGAGGAAAGATTGTGTTTATTAACCA 587
    |||||
```

Search completed: December 22, 2003, 23:11:34
Job time : 1114.05 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:59:40 ; Search time 37.4496 Seconds
(without alignments)
6682.695 Million cell updates/sec

Title: US-09-745-763-35_COPY_57_623
Perfect score: 567
Sequence: 1 CCTATCAGATTATCTTAACA.....AGATTGTTTATTAACCAA 567

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTCUTS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	567	100.0	1863	4	US-09-482-273-28 Sequence 28, Appl
2	49.2	8.7	7218	1	US-08-233-463-14 Sequence 14, Appl
3	33.6	5.9	1650	4	US-09-220-132-103 Sequence 103, Appl
4	33.2	5.9	2443	2	US-08-745-934-2 Sequence 2, Appl
5	33.2	5.9	9763	3	US-08-973-273-1 Sequence 1, Appl
6	33	5.8	1165	4	US-09-023-942A-28 Sequence 28, Appl
7	33	5.8	4791	4	US-09-328-352-1416 Sequence 1416, Ap
8	32.6	5.7	2489	3	US-09-141-047-7 Sequence 7, Appl
9	32.4	5.7	3483	4	US-09-620-312D-658 Sequence 658, App
10	32.2	5.7	17949	3	US-09-087-465-3 Sequence 3, Appl
11	31.8	5.6	615	4	US-09-107-532A-1080 Sequence 1080, Ap
12	31.8	5.6	2427	4	US-09-134-001C-268 Sequence 268, App
13	31.8	5.6	11811	3	US-09-078-294-7 Sequence 7, Appl
14	31.6	5.6	949	4	US-09-221-017B-594 Sequence 594, App
15	31.6	5.6	28473	4	US-08-961-527-83 Sequence 83, Appl
16	31.2	5.5	1712	3	US-08-809-336A-8 Sequence 8, Appl
17	31.2	5.5	1712	4	US-09-688-914A-8 Sequence 8, Appl
18	31.2	5.5	1712	4	US-09-688-913A-8 Sequence 8, Appl
19	31.2	5.5	1712	4	US-09-688-913A-8 Sequence 8, Appl
20	31.2	5.5	1927	4	US-09-689-916A-8 Sequence 8, Appl
21	31.2	5.5	1927	4	US-09-689-916A-13 Sequence 13, Appl
22	31.2	5.5	1927	4	US-09-689-916A-13 Sequence 13, Appl
23	31.2	5.5	1954	3	US-08-809-336A-13 Sequence 13, Appl
24	31	5.5	1155	4	US-09-134-001C-1364 Sequence 1364, Ap
25	31	5.5	1497	4	US-09-220-132-94 Sequence 94, Appl
26	31	5.5	1230025	4	US-09-198-452A-1 Sequence 1, Appl
27	30.8	5.4	707	3	US-08-998-416-1119 Sequence 1119, Ap

C	28	30.8	5.4	580073	4	US-08-545-528D-1	Sequence 1, Appl
	29	30.6	5.4	8000	4	US-09-415-784-101	Sequence 101, App
	30	30.6	5.4	8000	4	US-09-415-784-102	Sequence 102, App
	31	30.6	5.4	8000	4	US-09-415-785A-101	Sequence 101, App
	32	30.6	5.4	8000	4	US-09-415-785A-102	Sequence 102, App
	33	30.6	5.4	8000	4	US-08-944-465-101	Sequence 101, App
	34	30.6	5.4	8000	4	US-08-944-465-102	Sequence 102, App
	35	30.6	5.4	8000	4	US-09-415-868-101	Sequence 101, App
	36	30.6	5.4	8000	4	US-09-415-868-102	Sequence 102, App
	37	30.6	5.4	8000	4	US-09-415-900-101	Sequence 101, App
	38	30.6	5.4	8000	4	US-09-415-900-102	Sequence 102, App
	39	30.6	5.4	9951	4	US-09-193-707-3	Sequence 3, Appl
	40	30.6	5.4	10524	4	US-09-193-707-4	Sequence 4, Appl
	41	30.6	5.4	11282	4	US-09-733-042-1	Sequence 1, Appl
	42	30.6	5.4	11613	1	US-08-484-044-10	Sequence 10, Appl
	43	30.6	5.4	11703	1	US-08-801-263A-8	Sequence 8, Appl
	44	30.6	5.4	11703	3	US-09-102-248-8	Sequence 8, Appl
	45	30.6	5.4	11703	4	US-09-367-764-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-09-482-273-28
Sequence 28, Application US/09482273

Patent No. 6534631

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 71 Human Secreted Proteins

FILE REFERENCE: P2030P1

CURRENT APPLICATION NUMBER: US/09/482,273

CURRENT FILING DATE: 2000-01-13

EARLIER APPLICATION NUMBER: PCT/US99/15849

EARLIER FILING DATE: 1999-07-14

EARLIER APPLICATION NUMBER: 60/092,921

EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/092,922

EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/092,956

EARLIER FILING DATE: 1998-07-15

NUMBER OF SEQ ID NOS: 267

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 28

LENGTH: 1863

TYPE: DNA

ORGANISM: Homo sapiens

US-09-482-273-28

Query Match

Best Local Similarity 100.0%; Pred. No. 9.9e-166; Length 1863;

Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCTATCAGATTATCTTAACAAGAAACCAACTGGAAAAAATGAATTCCTATCTTC	60
DB	57	CCTATCAGATTATCTTAACAAGAAACCAACTGGAAAAAATGAATTCCTATCTTC	116
QY	61	GCATTTTCGGTGTGTTACCTTTATCCCTGTGCTTGGAAAGCTATATGCAAGAT	120
DB	117	GCATTTTCGGTGTGTTACCTTTATCCCTGTGCTTGGAAAGCTATATGCAAGAT	176
QY	121	GGCATCTCTAAGAGACCTTTTGAAGAAATATAAGATAGCAGCTGGAGATGT	180
DB	177	GGCATCTCTAAGAGACCTTTTGAAGAAATATAAGATAGCAGCTGGAGATGT	236
QY	181	GCTAAAGCAATCATCAACTAGCTGTATATGTAAGCCCAAGACAGATCTATAGCGA	240
DB	237	GCTAAAGCAATCATCAACTAGCTGTATATGTAAGCCCAAGACAGATCTATAGCGA	296
QY	241	TTGGACATCTTGGTGTGATCTGTTGACCCAGATGAGTGGCTCAAGAACTTGA	300
DB	297	TTGGACATCTTGGTGTGATCTGTTGACCCAGATGAGTGGCTCAAGAACTTGA	356

OY 301 GCATCCAAATATATACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCACTGAG 360
DB 357 GCATCCAAATATATATACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCACTGAG 416
OY 361 CCAAGTGAATATACCCCTGAGAGAGGAGAGAAATCACTGTGATCTGAGAGCAAGA 420
DB 417 CCAGTGAATATACCCCTGAGAGAGGAGAGAAATCACTGTGATCTGAGAGCAAGA 476
OY 421 ATTCAATAGATAGCATCTGGGCTTGTGAGAGAGCATTTGGGACCTCCAGAAAGCAT 480
DB 477 ATTCAATAGATAGCATCTGGGCTTGTGAGAGAGCATTTGGGACCTCCAGAAAGCAT 536
OY 481 ACAGAGAGATTTCTGTGATGACCTCTTTCATGATGACAGAGAGGAGCTCAGAGACA 540
DB 537 ACAGAGAGATTTCTGTGATGACCTCTTTCATGATGACAGAGAGGAGCTCAGAGACA 596
OY 541 AGAGGAGATTTGTTTATTAACCA 567
DB 597 AGAGGAGATTTGTTTATTAACCA 623

RESULT 2

US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOMULOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 8.7%; Score 49.2; DB 1; Length 7218;
Best Local Similarity 8.1%; Pred. No. 4,7e-05;
Matches 36; Conservative 215; Mismatches 193; Indels 0; Gaps 0;

OY 91 CTGTCTGTGGAAAGCTATATGCAAGATGCATCTCTAAGAGACCTTTGAAGATA 150
DB 1471 CTATGCAAGTATGTTAAAGATAGAGATTTGGTACRRRRRRRRRRRRRRRRRR 1412
OY 151 AAGAGAAATAGCACCTGTGAGATGTTGTAAGCAATCATCACTAGCTTTAT 210
DB 1411 RRR 1352
OY 211 GGTAAAGCCAGAAACATCTATGACGATTTGCCTTCTGTGATGATGAGACC 270
DB 1351 RRR 1292
OY 271 AGACTAGTGGCTCAGAACTAGAAAGCCATCAATATGATCAAACTGAC 330
DB 1291 RRR 1232
OY 331 CAAATGGGCTGAGAAAGTTCACTGAGCCAGTGAATACCCACTGGAGAGGGA 390
DB 1231 RRR 1172
OY 391 GAAGATCAGCTGTGATGCTGAGCCAGAAATTAATAGATGATCCTGCTTGGC 450
DB 1171 RRR 1112
OY 451 AGCAGCATGGGACTCTCCAGAGGATTAAGCAGAGATTTGTTGTGACCTTTG 510
DB 1111 RRR 1052
OY 511 GATGACTGAGAGAGGCTCA 534
DB 1051 GACCTGAGCAAGCTCGAATTA 1028

RESULT 3

US-09-220-132-103
Sequence 103, Application US/09220132
Patent No. 6506607

GENERAL INFORMATION:

APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
FILE REFERENCE: 07334-074001
CURRENT APPLICATION NUMBER: US/09/220,132
CURRENT FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 103

LENGTH: 1650

TYPE: DNA
ORGANISM: Homo sapiens
US-09-220-132-103

Query Match 5.9%; Score 33.6; DB 4; Length 1650;
Best Local Similarity 61.4%; Pred. No. 1.5;
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY 41 AATGAATTCCTATCTTGGCATTTTGGTGTGTCACCTTTATCCCTGTGCTG 100
DB 1313 AACTGAATGATTTGCTCCACATTAACGTTGGGCTGATTAATGCTCATGATG 1372
OY 101 GGAAGCTATATGCAAGATGCAATGCTC 128
DB 1373 AGAAGATATGCAAAATTAATATCTC 1400

RESULT 4

US-08-745-934-2/c
Sequence 2, Application US/08745934

Patent No. 5861496
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: HUMAN SQUALENE EPOXIDASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/745,934
FILING DATE: Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0151 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2443 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-08-745-934-2

Query Match 5.9%; Score 33.2; DB 2; Length 2443;
Best Local Similarity 50.0%; Pred. No. 2.5;
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 170 GTGAGATGTTGTTAAAGCAATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGAT 229
DB 741 GTATGTAAGTCCCGAAGCTTCTTATAAATAAGGTGAAGTGAATGCCAGAAAAGTC 682

QY 230 CCTATGAGCATTGGCACTTCTGTTATCTGTTGACCAACCACTGAGTGGCTCCAGA 289
DB 681 CACATGTTCCAAAGCTTCTTGTGTCGACCGGTCAAGCGGAGATTATCTGCTGAGTG 622

QY 290 ACCGAGAAAAGCCATCAATTATGTACCAAAACCTGCGACAGAGA 335
DB 621 GGTTAAGTTCTCCCAATGAGGTACCAAGAGACGAGAGA 576

RESULT 5
US-08-973-273-1/c
Sequence 1, Application US/08973273
Patent No. 6140085
GENERAL INFORMATION:
APPLICANT: Dean, Caroline
APPLICANT: MacKnight, Richard C
APPLICANT: Bancroft, Ian
APPLICANT: Lister, Clare K
TITLE OF INVENTION: Genetic Control of Flowering
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vandermhe P.C.

STREET: 1100 No. 6140085th Giebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,273
FILING DATE: 01-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/01332
FILING DATE: 03-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9511196.9
FILING DATE: 02-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ms Mary J Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-29
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9763 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
STRAIN: Columbia
US-08-973-273-1

Query Match 5.9%; Score 33.2; DB 3; Length 9763;
Best Local Similarity 61.6%; Pred. No. 5;
Matches 53; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 131 AGAGACTTTGGAAGAAATTAAGAGAAATATCCAGCTGTGAGATGTTGCTAAAGCA 190
DB 8640 AGAGAGTTCTGATCACTGATATAAACCACTACTGCTGTGAGATGTTGATATATCAT 8581

QY 191 TCATCAACTAGCTGTTATGTAA 216
DB 8580 TCATGAGAGAGCTCCACGCGGTATA 8555

RESULT 6
US-09-023-942A-28/c
Sequence 28, Application US/09023942A
Patent No. 6479274
GENERAL INFORMATION:
APPLICANT: (US only) ANTALIS Toni Marie and HOOBER John David
TITLE OF INVENTION: NOVEL MOLECULES
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,942A
; FILING DATE: 13-FEB-1998
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: POS101/97
; FILING DATE: 13-FEB-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PP0422/97
; FILING DATE: 18-NOV-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: International PCT Application
; FILING DATE: 13-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGICLO, FRANK S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 11168
; TELEPHONE: (516) 742 4363
; TELEFAX: (516) 742 4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-09-023-942A-28

Query Match
Best Local Similarity 5.8%; Score 33; DB 4; Length 1165;
Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Oy 280 GGCCTCAAGAACTTGAAGAAAGCCATCCAAATTATGTACCAAAACCTGCACCAAGATGG 339
Db 1126 GGTTCACAAATCTTAAATAATCATTAACATTATATACAAAAGCCGTGGGCGCCGG 1067
Oy 340 CTGAGAAAGTTCACTGAGCGGAGTGAATACCCCACTGGAGAGGGAGAAAGATCA 399
Db 1066 GCGGGGCGGAGTGTGGGGGCGGAGATGAGCGGTGGCGGGCGCTGATGCCCTTG 1007
Oy 400 GCTGTGATGTCGAGCC 416
Db 1006 GAGGCGGGGCTGAGGC 990

RESULT 7
US-09-328-352-1416
; Sequence 1416, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1416
; LENGTH: 4791
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-1416

Query Match
Best Local Similarity 5.8%; Score 33; DB 4; Length 4791;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Oy 124 ATCTCTAAGAGACTTTGAAAGAAATTAAGAAGAAATAGCCAGCTGTGAGAGTGTGCT 183
Db 1528 ATCATGTGGGTGTTTATGATGACAAAGCTGAGCATTAAGACGTTATGATAGATACC 1587
```

```

Oy 184 AAGCAATCATCAACCTAGCTGTTATGTATGAAGCCAGAGATCTTATGA 236
Db 1588 CAAGGAACTTAATTAAGCTATTGATCAGAAATGCCATACAAAGACTTATGA 1640

RESULT 8
US-09-141-047-7
; Sequence 7, Application US/09141047A
; Patent No. 6043085
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; TITLE OF INVENTION: Ehrlichia canis 120-kDa Immunodominant Antigenic
; FILE REFERENCE: D6143
; CURRENT APPLICATION NUMBER: US/09/141,047A
; CURRENT FILING DATE: 1998-08-27
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 7
; LENGTH: 2489
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; LOCATION: -340..2149
; OTHER INFORMATION: Nucleotide sequence of gene encoding 120kDa
; OTHER INFORMATION: Immunoreactive protein
; US-09-141-047-7

Query Match
Best Local Similarity 5.7%; Score 32.6; DB 3; Length 2489;
Matches 86; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Oy 15 TTACAGAAAGAAACCACTGGAAGAAAGAAATGAATTCCTATCTTCGATTTTCGGTGG 74
Db 538 TAAAGAAAGAAAGCTCTCTGAAGTTAAAGCGAAGATTTCACACCTGCTAGATGGTAG 597
Oy 75 TGTTCACTTTTATCCCTGTGCTCTGGAAAGCTATATGCAAGATGGCATCTTAAGAG 134
Db 598 TGTGAACATTCATCAAGTGAAGTGGGAAAGAAAGTATCGAAACTATGTAAGAGGAAAG 657
Oy 135 GACTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGAGATGTTGCTTAAGCA 189
Db 658 TACTCTGAAGTTAAAGCAGAAAGATTGCAACCTGCTGTAGATGGTAGTATAGAA 712

RESULT 9
US-09-620-312D-658
; Sequence 658, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Weinman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIORITY APPLICATION NUMBER: 09/552,317
```

PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pf_fl_genes Version 1.0
SEQ ID NO 658
LENGTH: 3483
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (94)..(2064)
US-09-620-312D-658

Query Match 5.7%; Score 32.4; DB 4; Length 3483;
Best Local Similarity 58.2%; Pred. No. 5.2;
Matches 57; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 101 GGAAGCTATATGCAAGATGCAATCTCTAAGAGGACTTTGAGAAATTAAGAGAA 160
DB 335 GTAAAGTTAGCCATCATGAGATTTGTAGCATTTTCAAGAGAGTGAAGAGAAC 394
QY 161 TAGCCAGCTGTGAGATGTTCCTAAGCAATCATCAAC 198
DB 395 TTGAAGCATTAAGGAGATGTTCAGCAATGAGCAAC 432

RESULT 10
US-09-087-465-3/c
Sequence 3, Application US/09087465A
Patent No. 6160092
GENERAL INFORMATION:
APPLICANT: Vinkemeier, Uwe
APPLICANT: Chen, Xiaomin
APPLICANT: Darnell Jr., James E
APPLICANT: Kuriyan, John
TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS OF
FILE REFERENCE: 600-1-229
CURRENT APPLICATION NUMBER: US/09/087,465A
CURRENT FILING DATE: 1998-05-29
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 17949
TYPE: DNA
ORGANISM: Homo sapiens
US-09-087-465-3

Query Match 5.7%; Score 32.2; DB 3; Length 17949;
Best Local Similarity 54.7%; Pred. No. 14;
Matches 64; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 5 TCAGATTATCTTAACAAGAAACCACTGAGAAAAAATTCCTTATCTTCGAT 64
DB 15045 TGAAGCTGTCTCAAAAAAATTTTCAAGAGTGAAGAGAAC 14986
QY 65 TTTTGGTGTGCTTATCTTATCTTCGCTCTGAGAAAGCTATATGCAAGATG 121
DB 14985 AGTGCCATGCTCATCTTCAATCCGCAATTTGGAGGCAAGAGAGACTG 14929

RESULT 11
US-09-107-532A-1080
Sequence 1080, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ardiello, Pamela Deneke

REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1080:

SEQUENCE CHARACTERISTICS:
LENGTH: 615 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...615
SEQUENCE DESCRIPTION: SEQ ID NO: 1080:
US-09-107-532A-1080

Query Match 5.6%; Score 31.8; DB 4; Length 615;
Best Local Similarity 49.7%; Pred. No. 3.3;
Matches 81; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 128 CTAAGAGACTTTTGAAGAAATAAAGAGAAATAGCCAGCTGTGAGATGTCCTAAG 187
DB 332 CTCAGGGAATCTTGAAAGAAAAACAAGATGATGAGAAATTAATTCGGCTGAA 391
QY 188 CAATCATCACTAGCTGTTTATGTTAAGCCAGAAAGATCTTATGAGGATGGCAC 247
DB 392 AAGCTAAAGAGAGCTCTGATTAAGATCCACTAAGAAAGCTTGGAAAAAGTTAGTGC 451
QY 248 TTTGCTGATACGTGTGAGCCAGACTGAGTGCCTCCAGAA 290
DB 452 TGAACATTGAAATGTTCTTTCGCAAGTATGCGGTACGAA 494

RESULT 12
US-09-134-001C-268
Sequence 268, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 268
LENGTH: 2427
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-268

Query Match
Best Local Similarity 5.6%; Score 31.8; DB 4; Length 2427;
Matches 69; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 30 ACTGAGAAAAAATGAATTCCTTATCTTGCATTTTGGTGCTGTCACCTTTATC 89
DB 2265 ACATGTAATAATATGTAATGATCATCTCGTAATTTCCGTGATGATGATTTAT 2324
QY 90 CCTGTCTCTGGGAAAGCTATATGCAAGATGGCATCTCTAAGAGACTTTGAAGAA 149
DB 2325 CGAGTCATCTGTAAGTATGCAAGTTAGCGGTATATCAATGATTAATTAAGATT 2384
QY 150 AAAAGAGAAA 160
DB 2385 AAAACAGAAA 2395

RESULT 13
US-09-078-294-7/c
Sequence 7, Application US/09078294
Patent No. 6265211
GENERAL INFORMATION:
APPLICANT: Choo, Kong-Hong Andy
APPLICANT: Du Sart, Desiree
APPLICANT: Cancilla, Michael R.
TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
FILE REFERENCE: Davies Col
CURRENT APPLICATION NUMBER: US/09/078,294
CURRENT FILING DATE: 1998-05-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 7
LENGTH: 11811
TYPE: DNA
ORGANISM: BAC-F2 config 3
US-09-078-294-7

Query Match
Best Local Similarity 5.6%; Score 31.8; DB 3; Length 11811;
Matches 60; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 16 TAACAGAAAAACCACTGGAAAAAATGAATTCCTTATCTTGCATTTTTCGGTGT 75
DB 6086 TAAAAAATAATGTAATTAATAAACAACCAACCAACCTGCTGTGATGTGGC 6027
QY 76 GTTCACTTTTATCCCTGTCTCTGGGAAAGCTATATGCAAGAAATG 122
DB 6026 TCACCCCTGTATCCCAACATTTGGAGGCGCAAGTGGATGATCG 5980

RESULT 14
US-09-221-017B-594
Sequence 594, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FORSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Montoy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 594:
SEQUENCE CHARACTERISTICS:
LENGTH: 949 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...949
US-09-221-017B-594

Query Match
Best Local Similarity 5.6%; Score 31.6; DB 4; Length 949;
Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 349 GTTCACTGGAGCCAGTGAGATACCCCACTGGGAGAGGGGAGAGATCAGTGTGATG 408
DB 251 GTACACGCGGCTGAATGCTATGACATCCCTGCTGACGCGTCAAGATACGCCGTTCTG 310
QY 409 CTGAGCCAGAGATTCATTAAGATAGC 434
DB 311 CTTGGCCATCATATAGAGAGAGC 336

RESULT 15
US-08-961-527-83/c
Sequence 83, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 22, 2003, 14:49:05 ; Search time 177.393 Seconds
(without alignments)
10650.710 Million cell updates/sec

Title: US-09-745-763-35_COPY_57_623

Perfect score: 567
Sequence: 1 CCTATCAGATTATCTTACA.....AGATTGTTGTTATACCAA 567

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2211978 seqs, 1666101734 residues

Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	567	100.0	1851	9	US-09-745-763-35
2	567	100.0	1863	11	US-09-984-271-28
3	399.6	70.5	1778	10	US-09-917-800A-505
4	247.4	43.6	427	10	US-09-833-981-1930
5	60	10.6	60	13	US-09-908-975-8960
6	41	7.2	63	13	US-09-908-975-1202
7	39.4	6.9	484	11	US-09-918-995-30889
8	39.4	6.9	1363	15	US-10-232-484-5
9	39.4	6.9	2319	14	US-10-044-090-579
10	39.4	6.9	3322	13	US-09-814-353-20277
11	39	6.9	665	13	US-10-027-632-101927
12	39	6.9	665	13	US-10-027-632-101927
13	36.6	6.5	786431	14	US-10-412-277-3
14	36.2	6.4	505	15	US-10-066-543-2072
15	36.2	6.4	1767	11	US-09-764-891-9681

C 16	36.2	6.4	715517	13	US-10-027-632-53712	Sequence 53712, A
C 17	36.2	6.4	715517	14	US-10-027-632-53712	Sequence 53712, A
C 18	36	6.3	17967	13	US-10-311-455-987	Sequence 987, App
C 19	35.2	6.2	513509	11	US-09-754-853A-4	Sequence 4, Appl1
C 20	35	6.2	390	11	US-09-918-995-35565	Sequence 35565, A
C 21	35	6.2	1136	13	US-10-027-632-249053	Sequence 249053, A
C 22	35	6.2	1136	14	US-10-027-632-249053	Sequence 249053, A
C 23	35	6.2	2048	13	US-10-094-749-1415	Sequence 1415, Ap
C 24	34.6	6.1	888	13	US-10-027-632-122347	Sequence 122347, A
C 25	34.6	6.1	888	14	US-10-027-632-122347	Sequence 122347, A
C 26	34.6	6.1	2155	13	US-10-027-632-252101	Sequence 252101, A
C 27	34.6	6.1	2155	14	US-10-027-632-252101	Sequence 252101, A
C 28	34.2	6.0	514	13	US-09-814-353-14775	Sequence 14775, A
C 29	34.2	6.0	552	13	US-09-814-353-2045	Sequence 2045, Ap
C 30	34.2	6.0	552	13	US-09-814-353-8391	Sequence 8391, Ap
C 31	34.2	6.0	708	13	US-10-027-632-28380	Sequence 28380, A
C 32	34.2	6.0	708	14	US-10-027-632-28380	Sequence 28380, A
C 33	34.2	6.0	5781	10	US-09-764-868-1374	Sequence 1374, Ap
C 34	34	6.0	783	13	US-10-027-632-62140	Sequence 62140, A
C 35	34	6.0	783	14	US-10-027-632-62140	Sequence 62140, A
C 36	34	6.0	882	9	US-09-815-242-4206	Sequence 4206, Ap
C 37	34	6.0	900	9	US-09-815-242-8477	Sequence 8477, Ap
C 38	34	6.0	5049	11	US-09-764-891-6588	Sequence 6588, Ap
C 39	34	6.0	5049	11	US-09-764-891-6589	Sequence 6589, Ap
C 40	34	6.0	5049	15	US-10-091-572-460	Sequence 460, App
C 41	34	6.0	5049	15	US-10-091-572-461	Sequence 461, App
C 42	34	6.0	34450	13	US-10-293-832-3	Sequence 3, Appl1
C 43	34	6.0	34757	13	US-10-293-832-22	Sequence 22, Appl1
C 44	33.6	5.9	789	13	US-10-027-632-123392	Sequence 123392, A
C 45	33.6	5.9	789	14	US-10-027-632-123392	Sequence 123392, A

ALIGNMENTS

RESULT 1
US-09-745-763-35
Sequence 35, Application US/09745763
Patent No. US20020065394A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaValle, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Merberg, David
Treacy, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 1851 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-745-763-35

Query Match 100.0%; Score 567; DB 9; Length 1851;
Best Local Similarity 100.0%; Pred. No. 9,2e-165;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCTATCGATTATCTTAAACAAGAAACCACTGGAAGAAATGAAATCTTATCTTC 60
DB 57 CCTATCGATTATCTTAAACAAGAAACCACTGGAAGAAATGAAATCTTATCTTC 116
QY 61 GCATTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
DB 117 GCATTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 176
QY 121 GGCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 180
DB 177 GGCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 236
QY 181 GCTTAAGCAATCATCAACCTGCTGTTATGTTAAAGCCAGAACATCTTATGAGCGA 240
DB 237 GCTTAAGCAATCATCAACCTGCTGTTATGTTAAAGCCAGAACATCTTATGAGCGA 296
QY 241 TTGGCACTTCTGTTGATCTGTTGAGACCCAGACTGAGTGCTCCAGAACTGAAAAA 300
DB 297 TTGGCACTTCTGTTGATCTGTTGAGACCCAGACTGAGTGCTCCAGAACTGAAAAA 356
QY 301 GCCATCCAAATTATGTACCAAAACCTGCGAGCAAGATGGCTGAGAAAGTTCACTGAG 360
DB 357 GCCATCCAAATTATGTACCAAAACCTGCGAGCAAGATGGCTGAGAAAGTTCACTGAG 416
QY 361 CCAGTGAATATACCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATCTGAGCCAAAG 420
DB 417 CCAGTGAATATACCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATCTGAGCCAAAG 476
QY 421 ATTCAATAGATAGCCATCTGGGTCTTTGGCAGACGATTTGGGACTCTCCAGAAAGCAT 480
DB 477 ATTCAATAGATAGCCATCTGGGTCTTTGGCAGACGATTTGGGACTCTCCAGAAAGCAT 536
QY 481 ACAGCAGAAAGTTCTGTGTGATGACCTTTTCGATGAATGCAAGAAAGGCTCAGAAACA 540
DB 537 ACAGCAGAAAGTTCTGTGTGATGACCTTTTCGATGAATGCAAGAAAGGCTCAGAAACA 596
QY 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
DB 597 AGAGGGAAGATTGTTGTTTATTAACCA 623
```

RESULT 2
US-09-984-271-28
Sequence 28, Application US/09984271
Publication No. US20030040088A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/984,271
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: 09/482,273
PRIOR FILING DATE: 2000-01-13
PRIOR APPLICATION NUMBER: PCT/US99/15849
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: 60/092,921
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/092,922
PRIOR FILING DATE: 1998-07-15

PRIOR APPLICATION NUMBER: 60/092,956
PRIOR FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 1863
TYPE: DNA
ORGANISM: Homo sapiens
US-09-984-271-28

Query Match 100.0%; Score 567; DB 11; Length 1863;
Best Local Similarity 100.0%; Pred. No. 9,2e-165;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCTATCGATTATCTTAAACAAGAAACCACTGGAAGAAATGAAATCTTATCTTC 60
DB 57 CCTATCGATTATCTTAAACAAGAAACCACTGGAAGAAATGAAATCTTATCTTC 116
QY 61 GCATTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
DB 117 GCATTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 176
QY 121 GGCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 180
DB 177 GGCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT 236
QY 181 GCTTAAGCAATCATCAACCTGCTGTTATGTTAAAGCCAGAACATCTTATGAGCGA 240
DB 237 GCTTAAGCAATCATCAACCTGCTGTTATGTTAAAGCCAGAACATCTTATGAGCGA 296
QY 241 TTGGCACTTCTGTTGATCTGTTGAGACCCAGACTGAGTGCTCCAGAACTGAAAAA 300
DB 297 TTGGCACTTCTGTTGATCTGTTGAGACCCAGACTGAGTGCTCCAGAACTGAAAAA 356
QY 301 GCCATCCAAATTATGTACCAAAACCTGCGAGCAAGATGGCTGAGAAAGTTCACTGAG 360
DB 357 GCCATCCAAATTATGTACCAAAACCTGCGAGCAAGATGGCTGAGAAAGTTCACTGAG 416
QY 361 CCAGTGAATATACCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATCTGAGCCAAAG 420
DB 417 CCAGTGAATATACCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATCTGAGCCAAAG 476
QY 421 ATTCAATAGATAGCCATCTGGGTCTTTGGCAGACGATTTGGGACTCTCCAGAAAGCAT 480
DB 477 ATTCAATAGATAGCCATCTGGGTCTTTGGCAGACGATTTGGGACTCTCCAGAAAGCAT 536
QY 481 ACAGCAGAAAGTTCTGTGTGATGACCTTTTCGATGAATGCAAGAAAGGCTCAGAAACA 540
DB 537 ACAGCAGAAAGTTCTGTGTGATGACCTTTTCGATGAATGCAAGAAAGGCTCAGAAACA 596
QY 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
DB 597 AGAGGGAAGATTGTTGTTTATTAACCA 623
```

RESULT 3
US-09-917-800A-505
Sequence 505, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Casale, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880

```

; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 505
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AF097723
US-09-917-800A-505

Query Match      70.5%; Score 399.6; DB 10; Length 1778;
Best Local Similarity 82.9%; Pred. No. 6,7e-113;
Matches 456; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 17 AACAGAAAAACAATGAGAAAAAATGAAATTCCTTATCTTGCAATTTTGGGTG 76
DB 95 AGCAGAGAAAAAGAACTAGAGGACATGAGTTCTTTCTTCCTGTTGCTGTG 154
QY 77 TTCAACCTTTATCCCTGCTCTGCGAAGCTATATGCAAGATGCGATCTTAAGGA 136
DB 155 TTCACTTTTCTCTGCGCTCTGGAAGAACTATATCAAGAGTGTTCACAGGAA 214
QY 137 CTTTGAAGAAATTAAGAAATAGCAGCTGTGAGATGTTGCTAAGACATCTCA 196
DB 215 CATTTCAAGAAATTAAGAAATAGCAGCTGTGAGATGTTGCTAAGACATCTCA 274
QY 197 ACCTAGCTTTTATGTTAAAGCCGAGAACAGATCTTAAGCAATTGGCACTTGTG 256
DB 275 ACCTGCTGTTTATGAAAAATACAGAACCGGTGTATGAGCGTTTGGACTTACTTG 334
QY 257 ATACTGTTGACCCAGCTAGTGGCTTCAAGAACTTAAGAAAAAGCATCAATTATGT 316
DB 335 ATACTGTTGACCCAGCTAGTGGCTTCAAGAACTTAAGAAAAAGCATCAATTATGT 394
QY 317 ACCAAAACTGACAGACATGAGGCTGAGAAAGTTCACTGAGCCAGTGAATACCC 376
DB 395 ACCAAAACTGACAGACATGAGGCTGAGAAAGTTCACTGAGCCAGTGAATACCC 454
QY 377 ACTGAGAGAGGAGAGAAATCAAGCTGTGATGCTGAGCCAGAAATTCATTAAGTACCA 436
DB 455 ACTGAGAGAGGAGAGAAATCTCAATGATGATGAGGCTCGAATTCACAAAGTTGCTA 514
QY 437 TCCTGGGCTTTGGCAGACATTTGGGCTCTCTCAAGAAAGCATTAACAGAGTTTGG 496
DB 515 TTTTAGGCTTTGGGAGCAGATTTGGGCTCTCTCAAGAAATTCACAGAGTACTCG 574
QY 497 TGGGACCTCTTTGATGAACTGAGAGAGGAGGCTCAGAAAGAGGAGGAGATTTTG 556
DB 575 TGGGAGCTCTTTTGTGTAACCTCAAGAGAGGAGATCAGAGAGGAGGAGATTTTG 634
QY 557 TTTATAACCA 566
DB 635 TTTATAACCA 644

RESULT 4
US-09-833-381-1930
; Sequence 1930, Application US/09833381
```

```

; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: PatSeq for Windows Version 3.0
; SEQ ID NO 1930
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(427)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1930

Query Match      43.6%; Score 247.4; DB 10; Length 427;
Best Local Similarity 87.3%; Pred. No. 4.4e-66;
Matches 269; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 75 TGTTCACCTTTATCCCTGCTGCTGGAAGCTATATGCAAGATGCGATCTTAAGAG 134
DB 117 TGTTCACCTTTATCCCTGCTGCTGGAAGCTATATGCAAGATGCGATCTTAAGAG 176
QY 135 GACTTTGAGAAATTAAGAAATAGCAGCTGTGAGATGTTGCTTAAGCAATCAT 194
DB 177 AACATTTGAGAAATTAAGAAATAGCAGCTGTGAGATGTTGCTTAAGCAATCAT 236
QY 195 CAACCTAGCTTTATGTTAAAGCCAGAACAGATCTTATGACGATTTGCTTCTGT 254
DB 237 CAACCTAGCTTTATGTTAAAGCCAGAACAGATCTTATGACGATTTGCTTCTGT 296
QY 255 TGATAGCTTTGAGCCAGACTGAGTGGCTCAAGAACTTAAGAAAAAGCATCAATTAT 314
DB 297 TGATAGCTTTGAGCCAGACTGAGTGGCTTAAGAACTTAAGAAAAAGCATCAATTAT 356
QY 315 GTACCAAACTGACAGCAAGATGGCTGAGAAAGTTCACTGAGCCAGTGAATAC 374
DB 357 GTACCAAACTGACAGCAAGATGGCTGAGAAAGTTCACTGAGCCAGTGAATAC 416
QY 375 CCACTGGG 382
DB 417 CCACTGGG 424

RESULT 5
US-09-908-975-8960
; Sequence 8960, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simcha
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 8960
; LENGTH: 60
```

TYPE: DNA
ORGANISM: Homo sapiens
US-09-908-975-8960

Query Match
Best Local Similarity 10.6%; Score 60; DB 13; Length 60;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

194 TCAACCTGCTGTTATGTTAAAGCCAGAACAGATCTTATGAGGATTGGCACTTGG 253
Db 1 TCAACCTGCTGTTATGTTAAAGCCAGAACAGATCTTATGAGGATTGGCACTTGG 60

RESULT 6
US-09-908-975-1202

Sequence 1202, Application US/09908975
Publication No. US20030165843A1

GENERAL INFORMATION:
APPLICANT: SHOSHAN, Avi
APPLICANT: WASSERMAN, Alon
APPLICANT: MINTZ, Eli
APPLICANT: FAIGLER, Simchon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
FILE REFERENCE: 36688-0005
CURRENT FILING DATE: 2001-07-20
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1202
LENGTH: 65
TYPE: DNA
ORGANISM: Rattus norvegicus
US-09-908-975-1202

Query Match

Best Local Similarity 7.2%; Score 41; DB 13; Length 65;
Matches 50; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

385 AGGGAGAGAAATCATGCTGATGCTGAGAGCCAGAAATTCATTAAGATCCATCTGGGT 444
Db 1 AGGGAGAGAAATCATGCTGATGCTGAGAGCCAGAAATTCATTAAGATCCATCTGGGT 60

445 CTTGG 449
61 CTTGG 65

RESULT 7

US-09-918-995-30889
Sequence 30889, Application US/09918995
Publication No. US20030073623A1

GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT FILING DATE: 2001-07-30
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: PastSeq for Windows Version 3.0
SEQ ID NO 30889
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: misc.feature
LOCATION: (1)..(484)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30889

Query Match
Best Local Similarity 6.9%; Score 39.4; DB 11; Length 484;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

258 TACTGTGACCCGACGATGAGTGTCTTCAAGAACTTGAAGAAAGCCATCCAAATTATGTA 317
Db 231 TAGGTCTGTTCCAAAGCACTGTGTCAAGAACTTGAAGAAATTGACATTAGGCAACAAC 290

318 CCAAACTGACGACAGATGAGGCTGAGAAAGTTTACCTGAGCCAGTGAATAACCCA 377
Db 291 CCTGAGAAATACACAGGGGACGCTTCCAGTGAATGTTGGGAGAGAGGAGGAG 350

378 CTGGAGAGGGAGAGAAATCAGCTGTGATGCTGAGCCAAAGAAATTCATTAAGATGACAT 437
Db 351 AGGGAGAGGGAGACAGATTCAGCTTGTGTGGGTCTGAGGGGTTCTTACAGGGGTAG 410

438 CTTGGGTTCTTGGCAGCGCATTTGGGACTC 466
Db 411 CCAAGATCTGGGAAACAGATCAGGACTC 439

RESULT 8
US-10-232-484-5

Sequence 5, Application US/10232484
Publication No. US20030113847A1

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Xonghong
APPLICANT: Drmanac, Radjye T
TITLE OF INVENTION: 79ICIP2BIV
FILE REFERENCE: 79ICIP2BIV
CURRENT FILING DATE: 2002-08-30
PRIOR FILING DATE: 2002-08-30
PRIOR APPLICATION NUMBER: 09/695,783
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: 09/552,929
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1363
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: CDS
LOCATION: (848)..(1090)
OTHER INFORMATION:
US-10-232-484-5

Query Match
Best Local Similarity 6.9%; Score 39.4; DB 15; Length 1363;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

258 TACTGTGACCCGACGATGAGTGTCTTCAAGAACTTGAAGAAAGCCATCCAAATTATGTA 317
Db 877 TAGGTCTGTTCCAAAGCACTGTGTCAAGAACTTGAAGAAATTGACATTAGGCAACAAC 936

318 CCAAACTGACGACAGATGAGGCTGAGAAAGTTTACCTGAGCCAGTGAATAACCCA 377
Db 937 CCTGAGAAATACCAAGGGGACGCTTCCAGTGAATGTTGGGAGAGAGAGGAG 996

378 CTGGAGAGGGAGAGAAATCAGCTGTGATGCTGAGCCAAAGAAATTCATTAAGATGACAT 437
Db 997 AGGGAGAGGGAGACAGATTCAGCTTGTGTGGGTCTGAGGGGTTCTTACAGGGGTAG 1056

438 CTTGGGTTCTTGGCAGCGCATTTGGGACTC 466
Db 1057 CCAAGATCTGGGAAACAGATCAGGACTC 1085


```
RESULT 9
US-10-044-090-579
; Sequence 579, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 579
; LENGTH: 2319
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 023244.1
; NAME/KEY: unsure
; LOCATION: 169,173
; OTHER INFORMATION: a, c, g, or other
US-10-044-090-579

Query Match          6.9%; Score 39.4; DB 14; Length 2319;
Best Local Similarity 49.3%; Pred. No. 0.32;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 258 TACTGTGGACCCAGACTGAGTGGCTCCAGAACCTTGAAGAAAGCCATCCAAATTATGTA 317
DB 1560 TAGTGTCTGTTCCAAAGCACTGTGTACAGAGCTTAATTCATTAGAGGACAAAC 1619
QY 318 CCAAAACCTGCAGCAAGATGGGCTGGAGAAAGTTCACTTGAGCCAGTGAATATCCCA 377
DB 1620 CCTGAGAAATACAAAGGGGCGACGCTTCCAGTAGATGTGTTGGGAGAGGAGGCGAG 1679
QY 378 CTGGAGAGGGGAGAGAGATGAGCTGTGATCTGAGCCAGCAAGAAATTCATAGATGACCAT 437
DB 1680 AGGGAGACAGGGGACAGATTCAGCTTGTGTGGTGGCTCTGAGGGTTCCTACAGGGGTAG 1739
QY 438 CCTGGGCTTTGGCAGCAGCATTTGGGACTC 466
DB 1740 CCAGATCTGGGAAACAGATCAGCGACTC 1768

RESULT 10
US-09-814-353-20277
; Sequence 20277, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
```

```
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20277
; LENGTH: 3322
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-20277

Query Match          6.9%; Score 39.4; DB 13; Length 3322;
Best Local Similarity 49.3%; Pred. No. 0.4;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 258 TACTGTGGACCCAGACTGAGTGGCTCCAGAACCTTGAAGAAAGCCATCCAAATTATGTA 317
DB 1946 TAGTGTCTGTTCCAAAGCACTGTGTACAGAGCTTAATTCATTAGAGGACAAAC 2005
QY 318 CCAAAACCTGCAGCAAGATGGGCTGGAGAAAGTTCACTTGAGCCAGTGAATATCCCA 377
DB 2006 CCTGAGAAATACAAAGGGGCGACGCTTCCAGTAGATGTGTTGGGAGAGGAGGCGAG 2065
QY 378 CTGGAGAGGGGAGAGAAATCAGCTGTGATCTGAGCCAGCAAGATTCATAGATGACCAT 437
DB 2066 AGGGAGACAGGGGACAGATTCAGCTTGTGTGGTGGCTCTGAGGGTTCCTACAGGGGTAG 2125
QY 438 CCTGGGCTTTGGCAGCAGCATTTGGGACTC 466
DB 2126 CCAGATCTGGGAAACAGATCAGCGACTC 2154

RESULT 11
US-10-027-632-101927
; Sequence 101927, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101927
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-101927

Query Match          6.9%; Score 39; DB 13; Length 665;
Best Local Similarity 63.2%; Pred. No. 0.21;
Matches 60; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 9 ATTATCTTAACAACAAACCACTGGAAGAAAAAATGGAATTCCTTATCTTCGATTGTTT 68
DB 389 AGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATGAGATGAGCTGGCTTTGTT 448
QY 69 CGGTGGGTTCACCTTTATCTCCCTGTGCTGGGA 103
DB 449 TGTGGCTCACACCTGTAATTCAGCACTTTGGGA 483
```

```

RESULT 12
US-10-027-632-101927
; Sequence 101927, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101927
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-101927

```

```

Query Match Similarity      6.9%; Score 39; DB 14; Length 665;
Best Local Similarity      63.2%; Pred. No. 0.21;
Matches      60; Conservative      0; Mismatches      35; Indels      0; Gaps      0

Oy      9      ATATCTTAAACAAGAAAACCAACTGGAAAAAATGAATTCCTTATCTTGCAATTTT      68
Db      389      AGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATGATGCTGGCTTGT      448

Oy      65      CGGTGCTGTTCACCTTTATCCCTGTCTCGGA      103
Db      445      TGGTGCTCACACTGTAAATTCACACACTTTGGGA      483

RESULT 13
US-10-412-277-3
; Sequence 3, Application US/10412277
; Publication No. US20030175791A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01067DIY
; CURRENT APPLICATION NUMBER: US/10/412,277
; CURRENT FILING DATE: 2003-04-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(786431)
; OTHER INFORMATION: n = A,T,C or G
US-10-412-277-3

Query Match      6.5%; Score 36.6; DB 13; Length 786431;
Best Local Similarity      50.3%; Pred.No. 64;

```

	Matches	90;	Conservative	0;	Mismatches	89;	Indels	0;	Gaps	0;
QY	13	TCTTRACAGAAAACCACTGGA	AAAAAAAAAATGAAATTCCTTATCTTCGCATTTTCGGT	72						
Db	537863	TCACAAATAGAAACAAATA	CAAGAAACCTAAATTTGTTCTTAAATAACTAATA	537922						
QY	73	GGTGTTCACCTTTATCCCTGTC	CTCTGGAAAGCATATGCAAGAAATGSCATCTCTAAG	132						
Db	537923	AAATTTCAAACTACTCTGACA	CAAAACAAATCAAGAAAAATATATAGGTATATAGTGAACAAAT	537982						
QY	133	AGCACTTTGAAAGAAATATAA	AGAAATATAGCCAGCTGTGAGATGTTGCTAAAGCAAT	191						
Db	537983	ATTACAGATGAAAAAGAAAA	AGAAATATATTAATCTGCTGATATAAATAATTTTAAAAATCAAT	538041						

```

RESULT 14
US-10-066-543-2072
; Sequence 2072, Application US/10066543
; Publication No. US20030087818a1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secretel, Heather
; APPLICANT: Cartier, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066.543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2072
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-2072

Query Match      6.4%; Score 36.2; DB 15; Length 505;
Best Local Similarity 60.8%; Pred. No. 1.3;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Cy      7 AGATTATCTTACACAGAAACCACTGGAAAAAATGAATTCCTATCTTGCAATT 66
      89 AAAATTAATTAATTAATGAAGAACAGATGTAATTAATTAATGAATTCCTGACCGGG 148
Cy      67 TTGGTGCTTACACTTTATCCCTGCTCTGGGA 103
      149 TGTGTGCTTACCGCTTAATATCCACAGACTTTGGGA 185
Db      149 TGTGTGCTTACCGCTTAATATCCACAGACTTTGGGA 185

RESULT 15
US-09-764-891-9681
; Sequence 9681, Application US/09764891
; Publication No. US20030077808a1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764.891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9681
; LENGTH: 1767
; TYPE: DNA

```

ORGANISM: Homo sapiens
US-09-764-891-9681

Query Match 6.4%; Score 36.2; DB 11; Length 1767;
Best Local Similarity 60.8%; Pred. No. 2.7;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY	7	AGATTATCTTAACAGAAACCACTGGAAAAAATGAATTCCTTATCTTCGCAATT	66
DB	481	AAATTAATTAATTAATGAAGAACAGATGAATATATTGAATGTGATTCCTGCGGG	540
QY	67	TTGGTGGTGTTCACCTTTTATCCCTGTGCTGGGA	103
DB	541	TGTGTGGCTTACGCGCTATATCCAGCACTTGGGA	577

Search completed: December 22, 2003, 23:44:46
Job time : 182.643 secs

THIS PAGE BLANK (USPTO)

QY 181 SRTVOYRTOGAVEAANKVGLASLIRSVASFSISPHHTGIOEYODSVPKIPACTITVEDAE 240
 DB 181 SRTVOYRTOGAVEAANKVGLASLIRSVASFSISPHHTGIOEYODSVPKIPACTITVEDAE 240
 QY 241 MMSRMAASHGKIIVYQKMGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVGGA 300
 DB 241 MMSRMAASHGKIIVYQKMGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVGGA 300
 QY 301 MDDGGGAFISWEALSLIKDLGRLPKRTLRVLMTAEBOGGVGAFOYQHLKVNISNYSLV 360
 DB 301 MDDGGGAFISWEALSLIKDLGRLPKRTLRVLMTAEBOGGVGAFOYQHLKVNISNYSLV 360
 QY 361 MMSDAGTFLPTGLOFTGSEKARAMEVWSLQPLNITQVLSHGEGTDINFWIQAGVGA 420
 DB 361 MMSDAGTFLPTGLOFTGSEKARAMEVWSLQPLNITQVLSHGEGTDINFWIQAGVGA 420
 QY 421 SLDDLYKXFFPHSHGDTMTVMDPKQNVAAVAVVSVVADMEMLPRS 472
 DB 421 SLDDLYKXFFPHSHGDTMTVMDPKQNVAAVAVVSVVADMEMLPRS 472
 RESULT 11
 ID AAY87081 standard; Protein; 473 AA.
 AC AAY87081;
 DT 09-MAY-2000 (first entry)
 DE Human secreted protein sequence SEQ ID NO:120.
 XX
 XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 XX antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;
 XX tumour; neurodegenerative disorder; developmental abnormality; allergy;
 XX foetal deficiency; blood disorder; immune system disorder; arthritis;
 XX autoimmune disease; hepatic disease; renal disease; inflammation;
 XX Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
 XX infection; AIDS; spinal cord injury; transplant rejection; diabetes;
 XX asthma; sepsis; acne; psoriasis; cardiovascular disorder;
 XX reproductive disorder; gastrointestinal disorder; respiratory disorder;
 XX metabolic disorder; food additive; preservative.
 OS Homo sapiens.
 PN WO200004140-A1.
 XX
 XX 27-JAN-2000.
 PD 14-JUL-1999; 99WO-US15849.
 PF 15-JUL-1998; 98US-0092921.
 PR 15-JUL-1998; 98US-0092922.
 PR 15-JUL-1998; 98US-0092956.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PB;
 PI Mucenki M, Endress GA, Soppet DR;
 XX WPI: 2000-161128/14.
 DR N-PSDB; AAZ98034.
 DR
 XX
 XX New isolated human genes, useful for diagnosis and treatment of, e.g.
 XX cancers, neurological or blood disorders
 XX
 XX Claim 11; Page 383-385; 494PP; English.
 XX
 XX The polynucleotide sequences given in AAZ98017 to AAZ98108 encode the
 CC human secreted proteins given in AAY87064 to AAY87223. Human secreted
 CC protein can have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;

CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
 CC antiallergic. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Human secreted protein s and their polynucleotides can
 CC be used for developing products for the diagnosis or treatment of cancer,
 CC tumours, neurodegenerative disorders, developmental abnormalities and
 CC foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation,
 CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The
 CC proteins or polynucleotides can also be used as food additives or
 CC preservatives. The proteins are also useful for identifying their
 CC binding partners. AAZ98008 to AAZ98016 and AAY87063 are sequence used in
 CC the exemplification of the present invention.
 CC
 SO Sequence 473 AA:
 Query Match 100.0%; Score 2417; DB 21; Length 473;
 Best Local Similarity 100.0%; Pred. No. 6.8e-212;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKFLIFAPFGVHLISLCSGKAICNGISKRTPEIKERINACGDVAKIINLVYGAQ 60
 DB 1 MKFLIFAPFGVHLISLCSGKAICNGISKRTPEIKERINACGDVAKIINLVYGAQ 60
 QY 61 NRSYERLALVDVTPRLSGSKNLEKAIQIMYONLQODGLEKVLHPRIPIHMERGESBA 120
 DB 61 NRSYERLALVDVTPRLSGSKNLEKAIQIMYONLQODGLEKVLHPRIPIHMERGESBA 120
 QY 121 VMLERIRIKIILIGGSSIGTPPEGITAEVLVVTSPFDELQRRASBARKIIVVQPYINY 180
 DB 121 VMLERIRIKIILIGGSSIGTPPEGITAEVLVVTSPFDELQRRASBARKIIVVQPYINY 180
 QY 181 SRTVOYRTOGAVEAANKVGLASLIRSVASFSISPHHTGIOEYODSVPKIPACTITVEDAE 240
 DB 181 SRTVOYRTOGAVEAANKVGLASLIRSVASFSISPHHTGIOEYODSVPKIPACTITVEDAE 240
 QY 241 MMSRMAASHGKIIVYQKMGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVGGA 300
 DB 241 MMSRMAASHGKIIVYQKMGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVGGA 300
 QY 301 MDDGGGAFISWEALSLIKDLGRLPKRTLRVLMTAEBOGGVGAFOYQHLKVNISNYSLV 360
 DB 301 MDDGGGAFISWEALSLIKDLGRLPKRTLRVLMTAEBOGGVGAFOYQHLKVNISNYSLV 360
 QY 361 MMSDAGTFLPTGLOFTGSEKARAMEVWSLQPLNITQVLSHGEGTDINFWIQAGVGA 420
 DB 361 MMSDAGTFLPTGLOFTGSEKARAMEVWSLQPLNITQVLSHGEGTDINFWIQAGVGA 420
 QY 421 SLDDLYKXFFPHSHGDTMTVMDPKQNVAAVAVVSVVADMEMLPRS 472
 DB 421 SLDDLYKXFFPHSHGDTMTVMDPKQNVAAVAVVSVVADMEMLPRS 472
 RESULT 12
 ID AAW33604 standard; Protein; 472 AA.
 AC AAW33604;
 DT 21-MAY-1998 (first entry)
 DE Human secreted protein AM282 full-length sequence.
 XX
 XX Secreted protein; AM282; cytokine; human.
 OS Homo sapiens.

THIS PAGE BLANK (USPTO)

```

XX Key Location/Qualifiers
FH Peptide 1..24
FT /label= Sig_peptide
FT Protein 25..472
FT /label= Mat_protein
PN WO9739030-A2.
XX
XX 23-OCT-1997.
XX
XX 16-APR-1997; 97WO-US06475.
XX
XX 13-JAN-1997; 97US-0783520.
XX 18-APR-1996; 96US-0634325.
XX
XX (GENY ) GENETICS INST INC.
XX
XX Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA,
XX Spaulding V;
XX
XX MPI: 1997-526400/48.
XX
XX N-PSDB; AAV02296.
XX
XX New isolated secretory proteins AM340, AM282 and AK583 - possibly
XX have cytokine, cell proliferation/differentiation regulating,
XX immunomodulating activities, etc.
XX
XX Claim 16; Page 45-47; 59pp; English.
XX
XX This human secreted protein, designated AM282, is encoded by a
XX full-length cDNA clone (see AAV02296), deposited in ATCC 98026, that
XX was identified from a database search using an isolated partial
XX AM282 clone (see AAV97398). AM282 protein can be used in a claimed
XX method for preventing, treating or ameliorating a medical
XX condition. It may exhibit cytokine, cell proliferation (either
XX inducing or inhibiting) or cell differentiation (either inducing or
XX inhibiting) activity or may induce production of other cytokines in
XX certain cell populations. It may also exhibit e.g. immune
XX stimulating or suppressing activity, haematopoiesis regulating
XX activity, tissue growth activity, activin/inhibin activity,
XX chemotactic or chemokinetic activity, haemostatic or thrombolytic
XX activity, receptor/ligand activity, anti-inflammatory activity,
XX tumour inhibition activity, or other activities. No evidence of
XX any of these activities is given in the specification.
XX
XX Sequence 472 AA;
SQ
Query Match 99.7%; Score 2410; DB 18; Length 472;
Best Local Similarity 99.8%; Pred. No. 3e-211;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

DB 301 MDDGGAFISWEALSLIKDLGRLPKRTLRVLTAEQGGVGAFOYQLHKVNISNYSLV 360
QY 361 MESDAGTFLPTGLQFTSEKRAIMEEYMSLQPLANTOVLSHEGTDINFIQAVPGA 420
DB 361 MESDAGTFLPTGLQFTSEKRAIMEEYMSLQPLANTOVLSHEGTDINFIQAVPGA 420
QY 421 SLDDLYKFFPFHSHGDTMTVMDPKQNNVAAVAVSVVADMEELPRS 472
DB 421 SLDDLYKFFPFHSHGDTMTVMDPKQNNVAAVAVSVVADMEELPRS 472

RESULT 13
AAB10229
ID AAB10229 standard; Protein, 472 AA.
XX
XX AAB10229;
XX
XX 16-NOV-2000 (first entry)
XX
XX Human fetal kidney protein fragment AM282_1i.
XX
XX Secreted protein; cytostatic; immunostimulatory; antimicrobial;
XX antiviral; immunosuppressive; antiinflammatory; vulnery; cytokine;
XX cell proliferation; differentiation; regulator; treatment; tumor;
XX autoimmune disease; inflammatory disorder; wound; microbial infection;
XX viral disease; graft versus host reaction suppression.
XX
XX Homo sapiens.
XX
XX WO200037630-A1.
XX
XX 29-JUN-2000.
XX
XX 22-DEC-1999; 99WO-US31005.
XX
XX 23-DEC-1998; 98US-0220876.
XX
XX (GENY ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C,
XX Merberg D, Treacy M, Bowman MR;
XX
XX MPI: 2000-442661/38.
XX N-PSDB; AAA40493.
XX
XX Secreted human proteins AS296-1i and AS34-1i, useful for treating
XX tumors, autoimmune diseases, inflammatory disorders, wounds, microbial
XX infections and viral diseases -
XX
XX Disclosure, Page 199-200; 293pp; English.
XX
XX This invention describes novel secreted human proteins (I) which have
XX cytostatic, immunostimulatory, antimicrobial, antiviral,
XX immunosuppressive, antiinflammatory and vulnery activity and which act
XX as cytokine, cell proliferation or differentiation regulators. (I)
XX is useful for treating tumors, autoimmune diseases, inflammatory
XX disorders, wounds, microbial infections and viral diseases. (I) is also
XX useful for suppressing graft versus host reaction. AAB10226-B10288
XX represent the secreted proteins encoded by AAA40490-A40580 which are
XX described in the method of the invention.
XX
XX Sequence 472 AA;
SQ
Query Match 99.7%; Score 2410; DB 21; Length 472;
Best Local Similarity 99.8%; Pred. No. 3e-211;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```


THIS PAGE BLANK (USPTO)

QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280
 DB 786 AAGACCTACCCAGATGACTGATTCCTTCAACACTGTACAGAGACTGGAGCAAAATAT 845
 QY 281 ProGluGlnValIleValSerGlyHisIleuAspSerTyrAspValGlyGlnGlyVal 300
 DB 846 CCGAGACAGGTTGACTGCTGACGTGACACTGTGACAGCTGGAGTGTGGCAGGGTCC 905
 QY 301 MetAspAspGlyGlyValAlaPheIleSerTyrGluIleuSerIleuIleuAspLeu 320
 DB 906 ATGGATATAGCCGGTGGAGCCTTATATATCATGGAGAGACTCTCACTTAATAAGACTT 965
 QY 321 GlyLeuArgProIleValArgThrIleuArgLeuValIleuTyrThrAlaGluGlnGlnGly 340
 DB 966 GGGGTGGCTCCAAAGAGAGACTCTCGCTGTGCTGTGACATGACAGAGAAACAAGTGA 1025
 QY 341 ValGlyAlaPheGlnTyrTyrGlnIleuHisIleValAsnIleSerAsnTyrSerIleuVal 360
 DB 1026 GTTGATGCTTCCTCGATTTATCACTTACACAGGTAATATTTCCAACTACAGTCTGGTG 1085
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
 DB 1086 ATGGAGTCTGACGACAGGAACCTTCTTACCCAGCTGGCTGCAATTCACCTGGCAGTGAAG 1145
 QY 381 AlaArgAlaIleMetGluGlnValIleMetSerIleuGlnProIleuAsnIleThrGlnVal 400
 DB 1146 GCCAGGGCCATCATGAGAGAGAGATATAGCTGTGACGCCCTCCAAATATCATCAGGTC 1205
 QY 401 LeuSerHisGlyGlyGlyThrAspIleAsnPheTyrIleGlnAlaGlyValProGlyVal 420
 DB 1206 CTGAGCCATGAG 1265
 QY 421 SerLeuLeuAspAspLeuTyrIleTyrPhePhePheHisHisSerHisGlyAspThrMet 440
 DB 1266 AGTCTACTGATGATCTTATACAGATATTTCTTCTCCATCCTCCACGAGAGACCATG 1325
 QY 441 ThrValMetAspProIleuGlnMetAsnValAlaIleAlaValIleThrAlaValIleSerTyr 460
 DB 1326 ACTGTCAATGATCCAAAGACATGAAAGTGTGCTGTGTTGGCTGTGTTGTTTAT 1385
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
 DB 1386 GTTGTTCAGACATGAGAGAAATGCTGCTAGTCC 1421

RESULT 13

AAV02296 ID AAV02296 standard; DNA; 1778 BP.

AC AAV02296;

DT 21-MAY-1998 (first entry)

DE Human secreted protein AM282 full-length cDNA clone.

KM Secreted protein; AM282; cytokine; human; db.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 17..1435

FT sig_peptide /+tag= a

FT mat_peptide /+tag= b

FT /+tag= b

XX WO9739030-A2.

XX 23-OCT-1997.

XX 16-APR-1997; 97MO-US06475.

XX 13-JAN-1997; 97US-0783520.

PR 18-APR-1996; 96US-0634325.
 XX (GENY) GENETICS INST INC.
 PA
 XX
 PI Jacobs K, LaValle ER, McCoy JM, Werberg D, Racie LA;
 PI Spaulding V;
 XX
 XX WPI, 1997-526400/48.
 DR P-PSDB; AAW3604.
 XX
 PT New isolated secretory proteins AM340, AM282 and AK583 - possibly
 PT have cytokine, cell proliferation/differentiation regulating,
 PT immunomodulating activities, etc.
 XX
 PS Claim 15; Page 44-45; 59pp; English.

CC This cDNA clone encodes a protein (see M33604) designated AM282.
 CC It was identified as "Yf95b10.r1 human EST 30142.5" (GenBank
 CC accession No. R77830) in a database search using a partial AM282
 CC clone (see T97398) obtained from a human foetal kidney cDNA
 CC library using methods selective for cDNAs encoding secreted
 CC proteins. AM282 is deposited in ATCC 98026 together with clones
 CC AM340 (see T97397) and AK583 (see V02297), which are also claimed.
 CC AM282 protein can be used in a claimed method for preventing,
 CC treating or ameliorating a medical condition. It may exhibit
 CC cytokine, cell proliferation (either inducing or inhibiting) or
 CC cell differentiation (either inducing or inhibiting) activity or
 CC may induce production of other cytokines in certain cell
 CC populations. It may also exhibit e.g. immune stimulating or
 CC suppressing activity, haematopoiesis regulating activity, tissue
 CC growth activity, activin/inhibin activity, chemotactic or
 CC chemokinetic activity, haemostatic or thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, tumour
 CC inhibition activity, or other activities. No evidence of any of
 CC these activities is given in the specification.

XX SQ - Sequence 1778 BP; 514 A; 386 C; 422 G; 456 T; 0 other;

Alignment Scores:

Pred. No.: 1,02e-235 Length: 1778
 Score: 2410.00 Matches: 471
 Percent Similarity: 99.79% Conservative: 0
 Best Local Similarity: 99.79% Mismatches: 1
 Query Match: 99.71% Indels: 0
 DB: 18 Gaps: 0

US-09-745-763-36 (1-472) x AAV02296 (1-1778)

QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20
 DB 17 ATGAATTCCTTATCTTCGCAATTTTCGGTGGTGTTCACCTTTATCCCTGTGCTGGG 76
 QY 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGlnIle 40
 DB 77 AAGACTTATGCAAGATGAGATCTTAAAGAGACTTTTAAAGAAATTAAGAAATA 136
 QY 41 AlaserCysGlyAspValAlaIleLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
 DB 137 GCCAGCTGTGAGAGTGTGCTTAAAGCAATCATCACTTCTGTATGTGTAAGCCAG 196
 QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
 DB 197 AACGATCCTATGAGCATTTGAGCACTTGTGATGATCTGTGAGACCCAGACTGAGTGC 256
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnGlnAspGlyLeu 100
 DB 257 TCCAGAACCTTAAGAAAGCAATCCAAATTTATGTACCAAACTGACGCAAGAGGCTG 316
 QY 101 GluLysValHisLeuGluProValArgIleProHisTyrProIleArgGlyGluGluSerAla 120
 DB 317 GAGAAAGTCACTCGAGCCAGTGAATATCCCACTGGGAGAGAGGAGAAAGATCAGCT 376
 QY 121 ValMetLeuGluProArgIleHisIleValIleLeuValLeuGluLysSerSerIleGly 140

THIS PAGE BLANK (USPTO)

```

DB      377  GTATGCTGAGCCAGAAATTCATAGATGCCCTCTGGCTTGCACAGCATGGG 436
QY      141  ThrProGluGlyIleThrAlaGluValIleuValThrSerPheAspGluLeuGln 160
DB      437  ACTCTCCAGAAAGCATTAACGAGAACTTCTGGTGTGACCTCTTTCGATGAACTGCAG 496
QY      161  ArgArgAlaSerGluAlaArgGlyLysIleValValIleValAsnGlnProTyrIleAsnTyr 180
DB      497  AGAAGGCGCTCAGAGAGCAGAGAGAGATTGTTGTTATACCACTTACATCACTAC 556
QY      181  SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaAlaIleValAlaLeu 200
DB      557  TCAAGACGCTGCAATACCAAGACGAGGGCGGTGGAGACCTGCACAGGCGGGCTTTG 616
QY      201  AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHISerThrGlyIleGln 220
DB      617  GCATCTCTCATTTGATCCGCGGCTCTTCTCCATCTACAGTCTCTCACAGGATTTGAG 676
QY      221  GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
DB      677  GAATACCAAGATGGCGGCCAGAAATTCACAGAGCTGTATTACGGTGGAAAGATGCAGAA 736
QY      241  MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
DB      737  ATGATGTCAAGATGGCTCTCTCATGGATCAAAATTTGTCATGAGCTAAAGATGGGGCA 796
QY      261  LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
DB      797  AAGACTTACCAGATCTGATCTTCTTACACACTGTACAGAGATCACTGGAGACCAATAT 856
QY      281  ProGluGlnValIleuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300
DB      857  CCAGAACAGGTGTACTAGTCACTGACATCTGACAGCTGGAGTGGGACGGATGCC 916
QY      301  MetAspAspGlyGlyGlyValAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
DB      917  ATGAGATGAGGGCGGTGAGCTTTATATATCAATGGAGACCTCTCATTTAAAGATCTT 976
QY      321  GlyLeuArgProLysArgThrLeuArgLeuValIleuTyrThrAlaGluGlnGlnGly 340
DB      977  GGGCTCGCTCCAAAGAGACTCTCGCGGTGGTCTGTGATCTGCAGAGAAACAGGTGGA 1036
QY      341  ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
DB      1037  GTTGGTGCCTTCAGATTATACATTACACAGGTAAATTTCCAACTACAGTCTGCTG 1096
QY      361  MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380
DB      1097  ATGGAGTCTGAGCGAGAACCTTCTTACCTGAGCTGCCTCAATTCACTGGCAGTGAAG 1156
QY      381  AlaArgAlaIleMetGluGlnValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
DB      1157  GCCAAGGCCATCATGAGAGAGATTATGACCTGCTGACCCCTCAATATCACTCAGGATC 1216
QY      401  LeuSerHisGlyGlyGlyThrAspIleAsnPheTyrIleGlnAlaGlyValProGlyAla 420
DB      1217  CTGAGCATGAGAGAGAGCAGACATCACTTTGATCCAAAGCTGAGATGCTGAGGCC 1276
QY      421  SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisSerHisGlyAspThrMet 440
DB      1277  AGTCACTTGAAGACTTATACAGATATTTCTTCTTCATCACTCCCAAGAGACACCAAG 1336
QY      441  ThrValMetAspProLysGlnMetAsnValAlaAlaValTyrAlaValIleSerTyr 460
DB      1337  ACTGCACTGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1396
QY      461  ValValAlaAspMetGluGlnMetLeuProArgSer 472
DB      1397  GTTGTTCAGACATGAGAAATGCTGCTTACGTTC 1432

```

RESULT 14
AAK94491

```

ID      AAK94491 standard; cDNA; 1860 BP.
XX
AC      AAK94491;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human full-length cDNA, SEQ ID NO: 3328.
XX
KW      Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS      Homo sapiens.
XX
PN      EP1130094-A2.
XX
PD      05-SEP-2001.
XX
PF      07-JUL-2000; 2000EP-0114089.
XX
PR      08-JUL-1999; 99JP-0194486.
PR      11-JAN-2000; 2000JP-0118774.
PR      02-MAY-2000; 2000JP-0183765.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y;
PI      Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR      WPI; 2001-524255/58.
XX
PT      P-PSDB; AAM93559.
XX
PT      830 Primers useful for synthesizing full length cDNA clones and their
PT      use in genetic manipulation -
XX
PS      Claim 8: SEQ ID NO 3328; 1380bp + sequence listing; English.
XX
CC      The invention relates to primers for synthesizing full length cDNA
CC      clones. 830 cDNA molecules encoding a human protein have been
CC      isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC      molecules have been determined. Primers for synthesizing the full length
CC      cDNA are useful for clarifying the function of the protein encoded by
CC      the cDNA. The full length clones were obtained by construction of full
CC      length enriched cDNA libraries that were synthesised by the oligo-capping
CC      method. The primers enable the production of the full length cDNA easily
CC      without any special methods. The present sequence is a full length
CC      human cDNA of the invention.
CC
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ      Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;

Alignment Scores:
Pred. No.: 7.14e-235 Length: 1860
Score: 2402.00 Matches: 469
Percent Similarity: 99.79% Conservative: 2
Best Local Similarity: 99.36% Mismatches: 1
Query Match: 99.38% Indels: 0
DB: 22 Gaps: 0

US-09-745-763-36 (1-472) x AAK94491 (1-1860)
QY      1  MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCysSerGly 20
DB      129  ATGAATATCTTATCTTGGCATTTTTCGTGTGTACCTTTATATCCCTGTGCTGGG 188
QY      21  LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
DB      189  AAAGCTATATCAAGAAAGATGATCTTCAAGAGACTTTTGAAGAAATGAAGAAATTA 248
QY      41  AlaSerCysGlyAspValAlaIleLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
DB      249  GCCAGCTGTGAGATGTTGCTAAAGCAATCAATCAACCTAGCTTTATGTAAGCCAG 308
QY      61  AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValIleGlyProArgLeuSerGly 80

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 22, 2003, 11:34:19 ; Search time 80.2473 Seconds
(without alignments)
933.601 Million cell updates/sec

Title: US-09-745-763-36
Perfect score: 2417
Sequence: 1 MKELFAFGVHLLSLICSG.....AVNAVSVVADMEMLPRS 472

Scoring table: BLASTM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq.19Jun03:*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	100.0	472	19	AAW85456
2	2417	100.0	472	21	AAW87254
3	2417	100.0	472	21	AAW58879
4	2417	100.0	472	22	AAW60658
5	2417	100.0	472	23	ABP61800
6	2417	100.0	472	23	ABG33880
7	2417	100.0	472	24	ABR47892
8	2417	100.0	472	24	ABR48106
9	2417	100.0	472	24	ABR00152

10	2417	100.0	472	24	ABR00274	Human gene 142 enc
11	2417	100.0	473	21	AAW87081	Human secreted pro
12	2410	99.7	472	18	AAW33604	Human secreted pro
13	2410	99.7	472	21	AAW10229	Human fetal kidney
14	2402	99.4	472	22	AAW93559	Human polypeptide
15	2378.5	98.4	461	22	AAW25762	Human protein sequ
16	2340	96.8	474	22	ABG09405	Novel human diagno
17	899	37.2	178	24	ABR48107	Human secreted pro
18	899	37.2	178	24	ABR00275	Human gene 142 enc
19	460	19.0	149	18	AAW27642	Secreted protein A
20	460	19.0	149	18	AAW31633	Human secreted pro
21	460	19.0	149	18	AAW44078	Human secreted pro
22	397	16.4	90	18	AAW84050	Human immune/haema
23	351	14.5	71	20	AAW11865	Human 5' EST secre
24	229.5	9.5	496	20	AAW89614	Aspergillus oryzae
25	229.5	9.5	496	20	AAW89597	Aspergillus oryzae
26	229.5	9.5	496	20	AAW89586	Aspergillus oryzae
27	226.5	9.4	481	21	AAW58020	Human peptidase NA
28	188	7.8	740	21	AAW58878	Human EST encoded
29	188	7.8	740	22	AAW24008	Human EST encoded
30	188	7.8	740	22	AAW88354	Human membrane or
31	173	7.2	750	21	AAW92644	Mutant human prost
32	169	7.0	333	17	AAW9416	Aminopeptidase pre
33	169	7.0	752	21	AAW92623	Murine prostate sp
34	169	7.0	761	21	AAW92660	Mutant murine pros
35	168	7.0	750	21	AAW92628	Mutant human prost
36	168	7.0	751	18	AAW31524	Prostate-specific
37	167.5	6.9	600	21	AAW21943	Arabidopsis thalia
38	167.5	6.9	610	21	AAW21942	Arabidopsis thalia
39	166	6.9	693	21	AAW92647	Mutant human PSM a
40	166	6.9	766	21	AAW92659	Mutant murine pros
41	165	6.8	33	21	AAW87185	Human secreted pro
42	165	6.8	33	22	AAW06125	Human gene 18 enco
43	165	6.8	33	22	AAW06162	Human gene 18 enco
44	165	6.8	33	23	ABG33947	Human secreted pro
45	165	6.8	33	23	ABG33985	Human secreted pro

ALIGNMENTS

RESULT 1	AAW85456	standard; Protein; 472 AA.
XX	AAW85456;	
AC	25-FEB-1999	(first entry)
XX		
DT	25-FEB-1999	(first entry)
XX		
DE	Secreted protein encoded by clone bu45_2.	
XX		
KW	Secreted protein; nutritional activity; immune stimulating; vaccine;	
KW	suppressing activity; haematopoiesis regulating activity;	
KW	tissue growth activity; activin; inhibin activity; chemotaxis;	
KW	chemokinetic activity; haemostasis; thrombolytic activity; receptor;	
KW	ligand; anti-inflammatory; cadherin; tumour invasion suppressor;	
KW	tumour inhibition; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
PN	W09842739-A2.	
XX		
PD	01-OCT-1998.	
XX		
PF	20-MAR-1998;	98WO-US05653.
XX		
PR	19-MAR-1998;	98US-0044466.
XX		
PR	21-MAR-1997;	97US-0822167.
XX		
PA	(GENY) GENETICS INST INC.	
XX		
PI	Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Werberg D;	
PI	Racle IA, Spaulding V, Treacy M;	

XX MPI, 1998-609890/51.
 DR N-PSDB; AAY82729.
 XX
 PT New polynucleotides encoding secreted human proteins - derived from
 PT human foetal brain, adult brain, foetal kidney, placenta or adult
 PT pineal gland cDNA libraries.
 XX
 PS Claim 14; Page 70-72; 113pp; English.
 XX
 CC The present sequence represents a secreted protein. The polynucleotide
 CC and secreted protein are predicted to have biological activities which
 CC would make them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is given.
 CC Suggested activities include nutritional activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cachectin/tumour
 CC invasion suppressor activity, and tumour inhibition activity (no data is
 CC given in the specification to support these activities). The
 CC polynucleotide is also stated to be useful for gene therapy.
 CC
 XX
 SQ Sequence 472 AA;
 Query Match 100.0%; Score 2417; DB 19; Length 472;
 Best Local Similarity 100.0%; Pred. No. 6,8e-212;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKFLIFAFPGVHLLSLCSGKAICNGISKRTFEEIKEEIASCGDVAKAIINLAVYKQAQ 60
 DB 1 MKFLIFAFPGVHLLSLCSGKAICNGISKRTFEEIKEEIASCGDVAKAIINLAVYKQAQ 60
 QY 61 NRSYERLALVDTVGPRLSGSKNLEKAIQIMYQNLQDGLKLVLEPRIPMERGESSA 120
 DB 61 NRSYERLALVDTVGPRLSGSKNLEKAIQIMYQNLQDGLKLVLEPRIPMERGESSA 120
 QY 121 VMLBRIHRIKIALIGSSIGTPEEGITAEVLVTSFDELORRASEARKIVVYNQPIYNY 180
 DB 121 VMLBRIHRIKIALIGSSIGTPEEGITAEVLVTSFDELORRASEARKIVVYNQPIYNY 180
 QY 181 SRTVOYRTQGAEEAAKAVGALASLRSVASFISYSPHTGISOEYDGVPKIPACTIVEDAE 240
 DB 181 SRTVOYRTQGAEEAAKAVGALASLRSVASFISYSPHTGISOEYDGVPKIPACTIVEDAE 240
 QY 241 MMSRAASHGIKIVITOLKNGAKTYPPTDSFNTVAETTSKRYPEQVVLVSGHLDSPVVGGA 300
 DB 241 MMSRAASHGIKIVITOLKNGAKTYPPTDSFNTVAETTSKRYPEQVVLVSGHLDSPVVGGA 300
 QY 301 MDDGGAFISWEALSLINDLGLRPRTLRVLVMTAEBOGVGAFQYOLAHKNISNYSVLV 360
 DB 301 MDDGGAFISWEALSLINDLGLRPRTLRVLVMTAEBOGVGAFQYOLAHKNISNYSVLV 360
 QY 361 MESDAGTFLPTGLOFTGSEKARAMEEWSLLOPINITQVLSHGEGTDINFIQAGVGA 420
 DB 361 MESDAGTFLPTGLOFTGSEKARAMEEWSLLOPINITQVLSHGEGTDINFIQAGVGA 420
 QY 421 SLDDLKYFFPFHSHGDTMTVMDBPKQNVAAAVVSYVADMBEMLPPS 472
 DB 421 SLDDLKYFFPFHSHGDTMTVMDBPKQNVAAAVVSYVADMBEMLPPS 472
 RESULT 2
 AAY87254
 ID AAY87254 standard; Protein; 472 AA.
 XX
 AC AAY87254;
 XX
 DT 11-MAY-2000 (first entry)
 XX
 DE Human signal peptide containing protein HSP-31 SEQ ID NO:31.
 XX
 KW Human; signal peptide-containing protein; HSP; diagnosis; cancer;

KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;
 KW antimicrobial; neuroprotective; cardiovascular; hepatocytic;
 KW antiaesthetic; gene therapy; cell proliferation; neurological disorder;
 KW reproductive disorder; developmental disorder; arteriosclerosis;
 KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
 KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
 KW Parkinson's disease; Huntington's disease; ovulatory defect;
 KW muscular dystrophy.
 KW
 OS Homo sapiens.
 XX
 XX WO200000610-A2.
 EN
 XX
 XX 06-JAN-2000.
 PD
 XX
 XX 25-JUN-1999; 99WO-US14484.
 PE
 XX
 XX 26-JUN-1998; 98US-0090762.
 PR
 XX 31-JUN-1998; 98US-0094983.
 PR 01-OCT-1998; 98US-0102686.
 PR 11-DEC-1998; 98US-0112129.
 PR
 XX
 PA (INCY-) INCYTE PHARM INC.
 PI Lal P, Tang YT, Gorgone GA, Corley NC, Guejler KJ, Baughn MR;
 PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
 PI Bandman O;
 XX
 XX WPI; 2000-160673/14.
 DR N-PSDB; AAZ98139.
 XX
 XX
 PT New human signal peptide-containing proteins useful in treatment,
 PT prevention and diagnosis of e.g. cancer, inflammation and
 PT cardiovascular disease -
 PT
 XX
 XX
 XX Claim 1; Page 179-180; 327pp; English.
 CC AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the
 CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have
 CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatocytic,
 CC neuroprotective, cardiovascular and antiaesthetic activities, and can
 CC be used in gene therapy. HSPs can be used to treat or prevent disorders
 CC associated with decreased activity or function of HSP. Antagonists of
 CC HSP are used to treat or prevent disorders associated with increased
 CC activity or function of HSP. Such disorders include cell proliferation
 CC (including cancer), inflammation, cardiovascular, neurological,
 CC reproductive or developmental disorders, (e.g. arteriosclerosis,
 CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
 CC asthma, Crohn's disease, microbial or other infections, congestive or
 CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
 CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP
 CC nucleic acids can be used for the recombinant production of HSP, for
 CC detecting HSP in standard hybridisation and amplification assays (for
 CC diagnosis and monitoring), in gene therapy, as antisense,
 CC triplex-forming or ribozyme therapeutics, for detecting related sequences
 CC or genetic variations, and for chromosomal mapping. HSP are also used to
 CC raise specific antibodies (Ab) and to screen for agonists and
 CC antagonists (potential therapeutic agents). Ab are used to diagnose, or
 CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic
 CC antagonists, in competitive drug screens, and for purification of HSP
 CC from natural sources.
 CC
 XX
 SQ Sequence 472 AA;
 Query Match 100.0%; Score 2417; DB 21; Length 472;
 Best Local Similarity 100.0%; Pred. No. 6,8e-212;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKFLIFAFPGVHLLSLCSGKAICNGISKRTFEEIKEEIASCGDVAKAIINLAVYKQAQ 60
 DB 1 MKFLIFAFPGVHLLSLCSGKAICNGISKRTFEEIKEEIASCGDVAKAIINLAVYKQAQ 60
 QY 61 NRSYERLALVDTVGPRLSGSKNLEKAIQIMYQNLQDGLKLVLEPRIPMERGESSA 120


```

Db      61 NRSYERIALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIIPMERGESBA 120
Qy      121 VMLPRHIAIILGLSSIGTPREGITAEVLVWVSFPELORRASEARGKIVVNOPIYNY 180
Db      121 VMLPRHIAIILGLSSIGTPREGITAEVLVWVSFPELORRASEARGKIVVNOPIYNY 180
Qy      181 SRTVQVTOGAVEAKGALASLIRSVASFSIYSPHTGISOYODGVPKIPRACTIVEDAE 240
Db      181 SRTVQVTOGAVEAKGALASLIRSVASFSIYSPHTGISOYODGVPKIPRACTIVEDAE 240
Qy      241 MMSHMAHGKIYIVQLKMGAKTYPDTSFNTVAEITGSKYPEQVVLVSGHLDMDVQGA 300
Db      241 MMSHMAHGKIYIVQLKMGAKTYPDTSFNTVAEITGSKYPEQVVLVSGHLDMDVQGA 300
Qy      301 MDDGGAFISWEALSLIKDGLRPRKTLRLVMTAEQGGVGARQYQLHKVNISNYSLV 360
Db      301 MDDGGAFISWEALSLIKDGLRPRKTLRLVMTAEQGGVGARQYQLHKVNISNYSLV 360
Qy      361 MESDAGTFLPTGLQFTGSEKARAIMBEVMSLLOPLINTQVLSHSGTDINFWIOAGVPGA 420
Db      361 MESDAGTFLPTGLQFTGSEKARAIMBEVMSLLOPLINTQVLSHSGTDINFWIOAGVPGA 420
Qy      421 SLDDLYKYFFPHSHSDTMTVMDPKQNNVAAVAVVSYVADMEEMLPSS 472
Db      421 SLDDLYKYFFPHSHSDTMTVMDPKQNNVAAVAVVSYVADMEEMLPSS 472

```

```

RESULT 3
AA58879
ID AA58879 standard; Protein; 472 AA.
AC AA58879;
XX
DT 08-MAY-2000 (first entry)
DE Human peptidase NALAD-ase IV.
XX
KW NALAD-ase IV; N-acetylated alpha-linked acidic dipeptidase; human;
KW chromosome 8q21.3; prostate cancer; neurodegenerative disease;
KW Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
KW peripheral neuropathy; Huntington's disease; acute brain injury;
KW multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;
KW therapy; diagnosis; neurotropic; neuroprotective; neuroleptic;
KW antiparkinsonian; anticonvulsant; vasotropic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 3..24
FT Modified-site 121 /note= "membrane-spanning domain"
FT Modified-site 179 /note= "N-glycosylated"
FT Modified-site /note= "N-glycosylated"
FT Modified-site 353 /note= "N-glycosylated"
FT Modified-site 356 /note= "N-glycosylated"
FT Modified-site /note= "N-glycosylated"
FT Modified-site 396 /note= "N-glycosylated"
XX
PN WO200004157-A2.
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-GB02241.
XX
PR 14-JUL-1999; 98GB-0015284.
XX
PA (JANNC) JANSSEN PHARM NV.
XX
PI Pangalos M, Neefs JEFM, Peeters DCG;

```

```

XX
DR WPI; 2000-182424/16.
DR N-PSDB; AA58813.
XX
PT New human N-acetylated alpha-linked acidic dipeptidases for treating
PT neural disorders e.g. Alzheimer's disease, schizophrenia and
PT Parkinson's disease
XX
PS Claim 15; Fig 5; 95pp; English.
XX
CC The present sequence is that predicted for human N-acetylated
CC alpha-linked acidic dipeptidase IV (NALAD-ase IV) on the basis of
CC isolated cDNA (see AA58813). NALAD-ase IV is predicted to be a
CC type II integral membrane protein of mol.wt. 51.9 kDa and pI 5.99.
CC Expression was low in all tissues examined by RT-PCR. The
CC invention provides human NALAD-ase I, II and IV cDNAs and encoded
CC polypeptides, as well as vectors, host cells, transgenic organisms,
CC antisense nucleic acids, agonists and antagonists. These are useful
CC for treating neural disorders such as Alzheimer's disease,
CC schizophrenia, ALS, Parkinson's disease, peripheral neuropathy,
CC Huntington's disease, acute brain injury, multiple sclerosis,
CC exposure to neurotoxins, peripheral nerve trauma, ischaemia or
CC dementia (claimed).
XX
SQ Sequence 472 AA;
XX
Query Match 100.0%; Score 2417; DB 21; Length 472;
Best Local Similarity 100.0%; Pred. No. 6.8e-212;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKFLIFAFGCVHLISCSGKALCKNGISKRTFEIEKEIASCGDVAKAIINLAVYKAQ 60
Db 1 MKFLIFAFGCVHLISCSGKALCKNGISKRTFEIEKEIASCGDVAKAIINLAVYKAQ 60
Qy 61 NRSYERIALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIIPMERGESBA 120
Db 61 NRSYERIALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIIPMERGESBA 120
Qy 121 VMLPRHIAIILGLSSIGTPREGITAEVLVWVSFPELORRASEARGKIVVNOPIYNY 180
Db 121 VMLPRHIAIILGLSSIGTPREGITAEVLVWVSFPELORRASEARGKIVVNOPIYNY 180
Qy 181 SRTVQVTOGAVEAKGALASLIRSVASFSIYSPHTGISOYODGVPKIPRACTIVEDAE 240
Db 181 SRTVQVTOGAVEAKGALASLIRSVASFSIYSPHTGISOYODGVPKIPRACTIVEDAE 240
Qy 241 MMSHMAHGKIYIVQLKMGAKTYPDTSFNTVAEITGSKYPEQVVLVSGHLDMDVQGA 300
Db 241 MMSHMAHGKIYIVQLKMGAKTYPDTSFNTVAEITGSKYPEQVVLVSGHLDMDVQGA 300
Qy 301 MDDGGAFISWEALSLIKDGLRPRKTLRLVMTAEQGGVGARQYQLHKVNISNYSLV 360
Db 301 MDDGGAFISWEALSLIKDGLRPRKTLRLVMTAEQGGVGARQYQLHKVNISNYSLV 360
Qy 361 MESDAGTFLPTGLQFTGSEKARAIMBEVMSLLOPLINTQVLSHSGTDINFWIOAGVPGA 420
Db 361 MESDAGTFLPTGLQFTGSEKARAIMBEVMSLLOPLINTQVLSHSGTDINFWIOAGVPGA 420
Qy 421 SLDDLYKYFFPHSHSDTMTVMDPKQNNVAAVAVVSYVADMEEMLPSS 472
Db 421 SLDDLYKYFFPHSHSDTMTVMDPKQNNVAAVAVVSYVADMEEMLPSS 472

```

```

RESULT 4
AAE06058
ID AAE06058 standard; Protein; 472 AA.
AC AAE06058;
XX
DT 24-SEP-2001 (first entry)
DE Human gene 18 encoded secreted protein HRACJ35, SEQ ID NO:120.
XX

```

Query Match	Similarity	Score	DB	Length	472:
Best Local	Similarity 100.0%	Pred. No. 6	8e-212		
Matches	472:	Conservative	0:	Mismatches	0:
				Indels	0:
				Gaps	0:
Qy	1	MKFLIFAFPGGVHLLSLCSGKALICKNGISKRTFEEIKEEIASCGDVAKAIINLAVGKAQ	60		
Db	1	MKFLIFAFPGGVHLLSLCSGKALICKNGISKRTFEEIKEEIASCGDVAKAIINLAVGKAQ	60		
Qy	61	NRSTVERALLVDVTYGPRLSGSKNLEKAIQIMYQMLQODGIEKXVLEFVRIPMERGESA	120		
Db	61	NRSTVERALLVDVTYGPRLSGSKNLEKAIQIMYQMLQODGIEKXVLEFVRIPMERGESA	120		
Qy	121	VMEPPRIHKIALILGLSSSIGTPPEGITAEVVLVYTSFDELQRRASEARGKITVYVQPIYNY	180		
Db	121	VMEPPRIHKIALILGLSSSIGTPPEGITAEVVLVYTSFDELQRRASEARGKITVYVQPIYNY	180		
Qy	181	SRIVQYRTQGAVEAKVAGALASLIRSVASFISYSPHTGICEYODGVKPIPTACTIVDEAE	240		
Db	181	SRIVQYRTQGAVEAKVAGALASLIRSVASFISYSPHTGICEYODGVKPIPTACTIVDEAE	240		
Qy	241	MMSTMAASHGKITIVQLMKAGAKTYPDTSFMTVAETISGSKPEQVVLVSGHLSMDVQGA	300		
Db	241	MMSTMAASHGKITIVQLMKAGAKTYPDTSFMTVAETISGSKPEQVVLVSGHLSMDVQGA	300		
Qy	301	MDDGGAGAFISWEALSLIKDLGRPKRTLRLVLMFAEEQGGVGAFQYQLHKVINSNLSLV	360		
Db	301	MDDGGAGAFISWEALSLIKDLGRPKRTLRLVLMFAEEQGGVGAFQYQLHKVINSNLSLV	360		
Qy	361	MESDAGTFLPTGILQFTSGSEKARAIMEEVMSLLOPLNTIYVLSHSGEPTDNFWIOAGVGA	420		
Db	361	MESDAGTFLPTGILQFTSGSEKARAIMEEVMSLLOPLNTIYVLSHSGEPTDNFWIOAGVGA	420		
Qy	421	SLLDLDLYKYPFEHSHGDTMTVMQKMNVAALAAVAVSVVVAEMELPRS	472		
Db	421	SLLDLDLYKYPFEHSHGDTMTVMQKMNVAALAAVAVSVVVAEMELPRS	472		
RESULT 5					
ABP61800					
ID ABP61800 standard; Protein; 472 AA.					
XX	AC	ABP61800;			
XX	DT	04-OCT-2002 (first entry)			
XX	DE	Human polypeptide SEQ ID NO 154.			
XX	XX	Human; cytosolic; antirheumatic; antiarthritic; vulnerary; analgesic;			
XX	XX	antiflammatory; antibacterial; immunosuppressive; antiparkinsonian;			
XX	XX	neuroprotective; nocotropic; osteopathic; haemostatic; vasotropic;			
XX	XX	antitumor; fungicide; antidiabetic; antiaslathmatic; antiallergic;			
XX	XX	immunostimulant; antiparasitic; secreted protein; transmembrane protein;			
XX	XX	cytokine; cell proliferation; cell differentiation; autoimmune disease;			
XX	XX	stem cell; growth factor; nervous system disease; neuropathy;			
XX	XX	Alzheimer's disease; Parkinson's disease; Huntington's disease;			
XX	XX	osteoporosis; severe combined immunodeficiency; SCID; infection;			
XX	XX	multiple sclerosis; rheumatoid arthritis; gene therapy.			
OS		Homo sapiens.			
XX	XX	US2002065394-A1.			
XX	XX	30-MAY-2002.			
XX	PD	22-DEC-2000; 2000US-0745763.			
XX	PR	18-MAR-1998; 98US-0040963.			
XX	PA	(JACO/) JACOBS K.			
XX	PA	(MCCO/) MCCOY J M.			
XX	PA	(LAVA/) LAVALLE E R.			
XX	PA	(COLL/) COLLINS-RACIE L A.			

PA (MERB/) MERBERG D.
 PA (TREAC/) TREACY M.
 PA (SPAUV/) SPAULDING V.
 XX
 PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 XX
 DR WPI: 2002-582343/62.
 DR N-PSDB: ABQ92016.
 XX
 PT Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis
 PT
 XX
 PS Claim 51; Page 114-115; 284pp; English.
 XX
 CC The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (CDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (I) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polypeptide of the invention.
 CC
 XX
 SQ Sequence 472 AA;
 Query Match 100.0%; Score 2417; DB 23; Length 472;
 Best Local Similarity 100.0%; Pred. No. 6.8e-212;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKPLIFAFGGVHLISLCSGKAIKNGISKPTPEIKKEIASCGVAAKAIINLAVYGAQ 60
 DB 1 MKPLIFAFGGVHLISLCSGKAIKNGISKPTPEIKKEIASCGVAAKAIINLAVYGAQ 60
 QY 61 NRSYERLALVDTVPRLSGSKNLEKAIQIMYONLQDGLKRVHLEPVRIPMERGESSA 120
 DB 61 NRSYERLALVDTVPRLSGSKNLEKAIQIMYONLQDGLKRVHLEPVRIPMERGESSA 120
 QY 121 VMLBPRHIAKILIGSSIGTPPEGITLAVLVTSFDELQRRASARAKIVVNPYINY 180
 DB 121 VMLBPRHIAKILIGSSIGTPPEGITLAVLVTSFDELQRRASARAKIVVNPYINY 180
 QY 181 SRTQYRTQGAVEAKVGALASLRVASFSISYHTGIOEODVPEKIPRACTIVEAE 240
 DB 181 SRTQYRTQGAVEAKVGALASLRVASFSISYHTGIOEODVPEKIPRACTIVEAE 240
 QY 241 MMSRMASHGIKIVIQLMKGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVQGA 300
 DB 241 MMSRMASHGIKIVIQLMKGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVQGA 300

DB 241 MMSRMASHGIKIVIQLMKGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVQGA 300
 QY 301 MDDGGGAFISWEALSLIKDGLRPKRTLRLVMTAERGQGVGAFOYVQLKVNINSNLSV 360
 DB 301 MDDGGGAFISWEALSLIKDGLRPKRTLRLVMTAERGQGVGAFOYVQLKVNINSNLSV 360
 QY 361 MESDAGTFLPTGLQFTGSEKARAIMERVMSLQPLNTQVLSHGEGTDINFWIQAGVPGA 420
 DB 361 MESDAGTFLPTGLQFTGSEKARAIMERVMSLQPLNTQVLSHGEGTDINFWIQAGVPGA 420
 QY 421 SLDDLKYTFPFHSHSDTMTMDPKOMNTAAAVMAVSVVADMEMLPRS 472
 DB 421 SLDDLKYTFPFHSHSDTMTMDPKOMNTAAAVMAVSVVADMEMLPRS 472
 RESULT 6
 ABG33880
 ID ABG33880 standard; Protein: 472 AA.
 XX
 AC ABG33880;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human secreted protein encoded by gene 18 #1.
 XX
 KW Human; secreted protein; gene therapy; immunosuppressive;
 KW antiarthritis; antirheumatic; antiproliferative; cyostatic; cardiant;
 KW vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;
 KW vinctide; fungicide; ophthalmological; autoimmune disease; neoplasm;
 KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
 KW ocular disorder; corneal infection; wound healing; skin aging;
 KW epithelial cell proliferation; food additive.
 KW
 XX
 OS Homo sapiens.
 OS
 XX
 PN WO200226931-A2.
 PD 04-APR-2002.
 PD
 XX
 PF 24-SEP-2001; 2001WO-US29871.
 PF
 XX
 PR 25-SEP-2000; 2000US-234925P.
 PR 12-JAN-2001; 2001WO-US00911.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y,
 PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Endress GA, Mucenaki M, Ebner R;
 DR
 DR WPI: 2002-362489/39.
 DR N-PSDB: ABK69743.
 XX
 PT Novel 71 isolated secreted polypeptides and polynucleotides encoding
 PT the polypeptides, useful for treating Huntington's disease, sepsis,
 PT meningitis, thrombocytopaenia, haemolytic anaemia, rheumatoid arthritis,
 PT asthma
 PT
 PS Claim 11; Page 1231-1232; 1478pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (or its
 CC fragment, homologue complement or allelic variant) encoding a human
 CC secreted protein (and its fragment, domain, epitope, variant, secreted
 CC form and species variant). Also included are a recombinant vector
 CC comprising the nucleic acid, a recombinant host cell comprising the
 CC vector, an antibody against the secreted protein, a recombinant host cell
 CC that expresses the secreted protein and a method of identifying a binding
 CC partner of the secreted protein. The nucleic acid and protein are used to
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
 CC for example autoimmune diseases e.g. rheumatoid arthritis,

hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. Many other diseases and disorders are listed in the specification. The polypeptides can also be used to aid wound healing in epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. The present sequence represents a novel human secreted protein of the invention.

Sequence 472 AA;

Query Match 100.0%; Score 2417; DB 23; Length 472;
Best Local Similarity 100.0%; Pred. No. 6.8e-212;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKFLIFAFPGVHLLSLCSGKAIQIMYQNLQDGLKVLKLEPVRIPIHMERGESB
1 NRSYERLALVDVTPGPRISGSKNLEKAIQIMYQNLQDGLKVLKLEPVRIPIHMERGESB
61 NRSYERLALVDVTPGPRISGSKNLEKAIQIMYQNLQDGLKVLKLEPVRIPIHMERGESB
121 VMLERIRHIAIILGLSSIGTPPEGITAEVAVVTSFDELORRASEARAGKIVVNOPIYNY
121 VMLERIRHIAIILGLSSIGTPPEGITAEVAVVTSFDELORRASEARAGKIVVNOPIYNY
181 SRTVOYRTOGAVEAKVAGLALSLRSVASFISYSHPTGIGIODEVPIPIACTIVEDAE
181 SRTVOYRTOGAVEAKVAGLALSLRSVASFISYSHPTGIGIODEVPIPIACTIVEDAE
241 MMSMAHSGIKIVIQKMGATYPTDTSFNTVAEITGSKYPEQVVLVSGHLDSDVVGGA
241 MMSMAHSGIKIVIQKMGATYPTDTSFNTVAEITGSKYPEQVVLVSGHLDSDVVGGA
301 MDDGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGVGAFQYQLHKVNISNYSLV
301 MDDGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGVGAFQYQLHKVNISNYSLV
361 MESDAGTFPLTGLPTGSEKARAIIMEEWSLLQPLNITQVLSHGEGTDINFWIQAGVGA
361 MESDAGTFPLTGLPTGSEKARAIIMEEWSLLQPLNITQVLSHGEGTDINFWIQAGVGA
421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVSYVADMEEMLPSS 472
421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVSYVADMEEMLPSS 472
421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVSYVADMEEMLPSS 472

RESULT 7
ABR47892
ID ABR47892 standard; Protein; 472 AA.

AC ABR47892;

DT 12-JUN-2003 (first entry)

DE Human secreted protein, SEQ ID 783.

XX Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytoskeletal;

XX vulnerable; antiinflammatory; neuroprotective;

XX antiparkinsonian; gene therapy; human; cardiovascular disorder.

OS Homo sapiens.

PN WO200295010-A2.

XX 28-NOV-2002.

XX 19-MAR-2002; 2002WO-US09785.

XX 21-MAR-2001; 2001US-277340P.
PR 19-JUN-2001; 2001US-306171P.
PR 13-NOV-2001; 2001US-331287P.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Ruben SM;
XX WPI; 2003-129429/12.

PT Novel human secreted proteins, useful for detecting, preventing,
PT diagnosing, prognosticating, treating and/or ameliorating
PT cardiovascular disorders such as arrhythmia.

PS Claim 13, SEQ ID 783; 1881bp; English.

CC The present invention relates to novel human secreted proteins
CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The
CC proteins and their coding sequences are useful for the preparation of a
CC diagnostic or pharmaceutical composition for diagnosing or treating a
CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
CC coronary arteriosclerosis and myocardial ischaemia), neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, renal disorders,
CC proliferative disorders and/or cancerous diseases and conditions, for
CC wound healing and epithelial cell proliferation, to treat inflammation or
CC infection, for treating thrombosis and arteriosclerosis, for treating or
CC preventing neural damage which occurs in neuronal disorders or
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
CC disease, to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts, to prevent skin aging or hair loss, to
CC stimulate growth and differentiation with other cytokines, to maintain
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism.

CC Note: The sequence data for this patent was published in electronic
CC format and is available from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 472 AA;

Query Match 100.0%; Score 2417; DB 24; Length 472;
Best Local Similarity 100.0%; Pred. No. 6.8e-212;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKFLIFAFPGVHLLSLCSGKAIQIMYQNLQDGLKVLKLEPVRIPIHMERGESB
1 NRSYERLALVDVTPGPRISGSKNLEKAIQIMYQNLQDGLKVLKLEPVRIPIHMERGESB
61 NRSYERLALVDVTPGPRISGSKNLEKAIQIMYQNLQDGLKVLKLEPVRIPIHMERGESB
121 VMLERIRHIAIILGLSSIGTPPEGITAEVAVVTSFDELORRASEARAGKIVVNOPIYNY
121 VMLERIRHIAIILGLSSIGTPPEGITAEVAVVTSFDELORRASEARAGKIVVNOPIYNY
181 SRTVOYRTOGAVEAKVAGLALSLRSVASFISYSHPTGIGIODEVPIPIACTIVEDAE
181 SRTVOYRTOGAVEAKVAGLALSLRSVASFISYSHPTGIGIODEVPIPIACTIVEDAE
241 MMSMAHSGIKIVIQKMGATYPTDTSFNTVAEITGSKYPEQVVLVSGHLDSDVVGGA
241 MMSMAHSGIKIVIQKMGATYPTDTSFNTVAEITGSKYPEQVVLVSGHLDSDVVGGA
301 MDDGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGVGAFQYQLHKVNISNYSLV
301 MDDGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGVGAFQYQLHKVNISNYSLV
361 MESDAGTFPLTGLPTGSEKARAIIMEEWSLLQPLNITQVLSHGEGTDINFWIQAGVGA
361 MESDAGTFPLTGLPTGSEKARAIIMEEWSLLQPLNITQVLSHGEGTDINFWIQAGVGA

PT New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 PT e.g. gastrointestinal diseases and disorders, or cancers -

PS Claim 13; Page 1029-1030; 1216pp; English.

CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein of the
 CC invention.

XX Sequence 472 AA;

Query Match 100.0%; Score 2417; DB 24; Length 472;
 Best Local Similarity 100.0%; Pred. No. 6.8e-212;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFGGVHLLSLCSGKAICKNGISKRTPEEIKBEIASCGVAKAIINLAAYGKAQ 60
 DB 1 MKFLIFAFGGVHLLSLCSGKAICKNGISKRTPEEIKBEIASCGVAKAIINLAAYGKAQ 60
 QY 61 NRSYERLALVDYGPRLSSGKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120
 DB 61 NRSYERLALVDYGPRLSSGKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120
 QY 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPFDELORASBARSKIVVYNQPYINY 180
 DB 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPFDELORASBARSKIVVYNQPYINY 180
 QY 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPFDELORASBARSKIVVYNQPYINY 180
 DB 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPFDELORASBARSKIVVYNQPYINY 180
 QY 181 SRTVOYRTOGAVEAKVAGALASLISVASFISYSPHTGIOEYODGVPRIPACTIVEDAE 240
 DB 181 SRTVOYRTOGAVEAKVAGALASLISVASFISYSPHTGIOEYODGVPRIPACTIVEDAE 240
 QY 181 SRTVOYRTOGAVEAKVAGALASLISVASFISYSPHTGIOEYODGVPRIPACTIVEDAE 240
 DB 181 SRTVOYRTOGAVEAKVAGALASLISVASFISYSPHTGIOEYODGVPRIPACTIVEDAE 240
 QY 241 MMSRASHGKITIVIQKMGAKTYPPTDSFNTVAETITGSKYPRQVVLVSGHLDMSWVGGA 300
 DB 241 MMSRASHGKITIVIQKMGAKTYPPTDSFNTVAETITGSKYPRQVVLVSGHLDMSWVGGA 300
 QY 241 MMSRASHGKITIVIQKMGAKTYPPTDSFNTVAETITGSKYPRQVVLVSGHLDMSWVGGA 300
 DB 241 MMSRASHGKITIVIQKMGAKTYPPTDSFNTVAETITGSKYPRQVVLVSGHLDMSWVGGA 300
 QY 301 MDDGGAFISWEALSLINDLGLRPRKTRLVMTAEEOGVGAFQYQOLHKVNISNYSLV 360
 DB 301 MDDGGAFISWEALSLINDLGLRPRKTRLVMTAEEOGVGAFQYQOLHKVNISNYSLV 360
 QY 301 MDDGGAFISWEALSLINDLGLRPRKTRLVMTAEEOGVGAFQYQOLHKVNISNYSLV 360
 DB 301 MDDGGAFISWEALSLINDLGLRPRKTRLVMTAEEOGVGAFQYQOLHKVNISNYSLV 360
 QY 361 MESDAGTLPGLTGTGSEKARAIIEEYWSLIQPINITQVLSHGEGTDINFWIQAGVGA 420
 DB 361 MESDAGTLPGLTGTGSEKARAIIEEYWSLIQPINITQVLSHGEGTDINFWIQAGVGA 420
 QY 421 SLDDLYKFFPFHSHGDTMTVMQKNVAAVAVVSYVVADEEMLPSS 472
 DB 421 SLDDLYKFFPFHSHGDTMTVMQKNVAAVAVVSYVVADEEMLPSS 472

RESULT 10
 ABR00274
 ID ABR00274 standard; Protein; 472 AA.

XX ABR00274;
 XX
 DT 03-APR-2003 (first entry)

XX Human gene 142 encoded secreted protein HRACJ35, SEQ ID NO:563.

DE Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
 KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
 KW antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2.

OS Homo sapiens.

PN WO200276488-A1.

XX 03-OCT-2002.

XX 19-MAR-2002; 2002WO-US08276.

XX 21-MAR-2001; 2001US-277340P.

XX 19-JUN-2001; 2001US-306171P.

XX 13-NOV-2001; 2001US-331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-029900/02.

XX N-PSDB; AB271453.

PT New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 PT e.g. gastrointestinal diseases and disorders, or cancers -

PS Claim 13; Page 1108-1109; 1216pp; English.

CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein of the
 CC invention.

XX Sequence 472 AA;

Query Match 100.0%; Score 2417; DB 24; Length 472;
 Best Local Similarity 100.0%; Pred. No. 6.8e-212;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFGGVHLLSLCSGKAICKNGISKRTPEEIKBEIASCGVAKAIINLAAYGKAQ 60
 DB 1 MKFLIFAFGGVHLLSLCSGKAICKNGISKRTPEEIKBEIASCGVAKAIINLAAYGKAQ 60
 QY 61 NRSYERLALVDYGPRLSSGKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120
 DB 61 NRSYERLALVDYGPRLSSGKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120
 QY 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPFDELORASBARSKIVVYNQPYINY 180
 DB 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPFDELORASBARSKIVVYNQPYINY 180

QY 181 SRTVYRTQGAWEAKVAGALASLIRSVASFSIYSPHTGIOEYODGVPIPTACTIVEDAE 240
DB 181 SRTVYRTQGAWEAKVAGALASLIRSVASFSIYSPHTGIOEYODGVPIPTACTIVEDAE 240
QY 241 MMSRMASHGKIYIVQLKMGAKATYPTDTSFNTVAEITGSKYPEQVVLVSGHLDSDMDVGGA 300
DB 241 MMSRMASHGKIYIVQLKMGAKATYPTDTSFNTVAEITGSKYPEQVVLVSGHLDSDMDVGGA 300
QY 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEBOGGVGAPOYYQLHKVNISNTSLV 360
DB 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEBOGGVGAPOYYQLHKVNISNTSLV 360
QY 361 MESDAGTFLPTGLQFTGSEKARAIEMEYMSLQPLNTIQTUOVLSHGEGTDINFWIQAQVPGA 420
DB 361 MESDAGTFLPTGLQFTGSEKARAIEMEYMSLQPLNTIQTUOVLSHGEGTDINFWIQAQVPGA 420
QY 421 SLDDLKYKFPFHSHGDTMTVMDPKQNNVAAVAVSVYVADMEMLPRS 472
DB 421 SLDDLKYKFPFHSHGDTMTVMDPKQNNVAAVAVSVYVADMEMLPRS 472

RESULT 11

AA87081
ID AA87081 standard; Protein; 473 AA.
XX
XX AA87081;
XX
XX 09-MAY-2000 (first entry)
XX
XX Human secreted protein sequence SEQ ID NO:120.
XX

Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
antiinflammatory; neurotropic; neuroprotective; antiallergic; cancer;
tumour; neurodegenerative disorder; developmental abnormality; allergy;
foetal deficiency; blood disorder; immune system disorder; arthritis;
autoimmune disease; hepatic disease; renal disease; inflammation;
Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
infection; AIDS; spinal cord injury; transplant rejection; diabetes;
asthma; sepsis; acne; psoriasis; cardiovascular disorder;
reproductive disorder; gastrointestinal disorder; respiratory disorder;
metabolic disorder; food additive; preservative.

Homo sapiens.

OS
XX
XX WO200004140-A1.
XX
XX 27-JAN-2000.
XX
XX PD
XX PF 14-JUL-1999; 99WO-US15849.
XX
XX PR 15-JUL-1998; 98US-0092921.
XX PR 15-JUL-1998; 98US-0092922.
XX PR 15-JUL-1998; 98US-0092956.

(HUMA-) HUMAN GENOME SCI INC.

XX
XX Ruben SM, Komatoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;
PI Mucenski M, Endress GA, Soppet DR;
XX
XX MPI: 2000-161128/14.
XX
XX N-PSDB; AA298034.

XX
XX New isolated human genes, useful for diagnosis and treatment of, e.g.
XX
XX cancer, neurological or blood disorders

XX
XX Claim 11; Page 383-385; 494pp; English.

XX
XX The polynucleotide sequences given in AA298017 to AA298108 encode the
XX
XX human secreted proteins given in AA87064 to AA87223. Human secreted
XX
XX protein can have activities based on the tissues and cells the genes are
XX
XX expressed in. Examples of activities include: cytostatic;

CC immunosuppressive; antiinflammatory; neurotropic; neuroprotective; and
CC antiallergic. The polynucleotides and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Human secreted protein s and their polynucleotides can
CC be used for developing products for the diagnosis or treatment of cancer,
CC tumours, neurodegenerative disorders, developmental abnormalities and
CC foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The
CC proteins or polynucleotides can also be used as food additives or
CC preservatives. The proteins are also useful for identifying their
CC binding partners. AA298008 to AA298016 and AA87063 are sequence used in
CC the exemplification of the present invention.

Sequence 473 AA:

Query Match 100.0%; Score 2417; DB 21; Length 473;
Best Local Similarity 100.0%; Pred. No. 6,8e-212;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAPFGVHLISCSGKATIKNGISKTEFEIEEETASCGDVAKAIINLAVYKAQ 60
DB 1 MKFLIFAPFGVHLISCSGKATIKNGISKTEFEIEEETASCGDVAKAIINLAVYKAQ 60
QY 61 NRSYERLALVDYTPGLSGSKNLEKAIQIOMYQULQODGLEKVLAEVRIIPMERGESA 120
DB 61 NRSYERLALVDYTPGLSGSKNLEKAIQIOMYQULQODGLEKVLAEVRIIPMERGESA 120
QY 121 VMLBPRILHKAIIIGLSSIGTPPEGITAEVLVNTSPDELQRASBARKIVVYQPYINY 180
DB 121 VMLBPRILHKAIIIGLSSIGTPPEGITAEVLVNTSPDELQRASBARKIVVYQPYINY 180
QY 181 SRTVYRTQGAWEAKVAGALASLIRSVASFSIYSPHTGIOEYODGVPIPTACTIVEDAE 240
DB 181 SRTVYRTQGAWEAKVAGALASLIRSVASFSIYSPHTGIOEYODGVPIPTACTIVEDAE 240
QY 241 MMSRMASHGKIYIVQLKMGAKATYPTDTSFNTVAEITGSKYPEQVVLVSGHLDSDMDVGGA 300
DB 241 MMSRMASHGKIYIVQLKMGAKATYPTDTSFNTVAEITGSKYPEQVVLVSGHLDSDMDVGGA 300
QY 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEBOGGVGAPOYYQLHKVNISNTSLV 360
DB 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEBOGGVGAPOYYQLHKVNISNTSLV 360
QY 361 MESDAGTFLPTGLQFTGSEKARAIEMEYMSLQPLNTIQTUOVLSHGEGTDINFWIQAQVPGA 420
DB 361 MESDAGTFLPTGLQFTGSEKARAIEMEYMSLQPLNTIQTUOVLSHGEGTDINFWIQAQVPGA 420
QY 421 SLDDLKYKFPFHSHGDTMTVMDPKQNNVAAVAVSVYVADMEMLPRS 472
DB 421 SLDDLKYKFPFHSHGDTMTVMDPKQNNVAAVAVSVYVADMEMLPRS 472

RESULT 12

AA87081
ID AA87081 standard; Protein; 472 AA.

XX
XX AA87081;
XX
XX 21-MAY-1998 (first entry)
XX
XX Human secreted protein AM282 full-length sequence.

XX
XX Secreted protein; cytokine; human.
XX
XX Homo sapiens.


```

XX Key Location/Qualifiers
FH Peptide 1..24
FT /label= Sig_peptide
FT Protein 25..472
FT /label= Mat_protein
XX WO9739030-A2.
XX 23-OCT-1997.
XX 16-APR-1997; 97WO-US06475.
XX 13-JAN-1997; 97US-0783520.
XX 18-APR-1996; 96US-0634325.
XX (GENY ) GENETICS INST INC.
XX Jacobs K, LaVallie ER, McCoy JM, Merberg D, Racie LA;
PI Spaulding V;
XX MPI: 1997-526400/48.
XX N-PSDB; AAV02296.
XX New isolated secretory proteins AM340, AM282 and AK583 - possibly
PT have cytokine, cell proliferation/differentiation regulating,
PT immunomodulating activities, etc.
XX Claim 16; Page 45-47; 59pp; English.
XX This human secreted protein, designated AM282, is encoded by a
CC full-length cDNA clone (see AAV02296), deposited in ATCC 98026, that
CC was identified from a database search using an isolated partial
CC AM282 clone (see AAT97398). AM282 protein can be used in a claimed
CC method for preventing, treating or ameliorating a medical
CC condition. It may exhibit cytokine, cell proliferation (either
CC inducing or inhibiting) or cell differentiation (either inducing or
CC inhibiting) activity or may induce production of other cytokines in
CC certain cell populations. It may also exhibit e.g. immune
CC stimulating or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic or chemokinetic activity, haemostatic or thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC tumour inhibition activity, or other activities. No evidence of
CC any of these activities is given in the specification.
XX
SQ Sequence 472 AA;
Query Match 99.7%; Score 2410; DB 18; Length 472;
Best Local Similarity 99.8%; Pred. No. 3e-211;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKFLIPAFGGVHLLSLCSGKAIICKNGISKRTPEEIKKEIASCGDVAKIINLVYGRQA 60
DB 1 MKFLIPAFGGVHLLSLCSGKAIICKNGISKRTPEEIKKEIASCGDVAKIINLVYGRQA 60
QY 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYQNLQDDLEKVLHPVRIIPMERGESBA 120
DB 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYQNLQDDLEKVLHPVRIIPMERGESBA 120
QY 121 VMLERIRIKHAIILGLSSIGTPPEGITAEVLVVTSTDELRASERGRKIVVYNGPYNY 180
DB 121 VMLERIRIKHAIILGLSSIGTPPEGITAEVLVVTSTDELRASERGRKIVVYNGPYNY 180
QY 181 SFTVOYRTGAAVEAAKVGALASLIRSVASFISYSPHTGIGIYQDGVKIPITACTIVEDAE 240
DB 181 SFTVOYRTGAAVEAAKVGALASLIRSVASFISYSPHTGIGIYQDGVKIPITACTIVEDAE 240
QY 241 MMSRWASHSIKIVIQIKMGAKTYPTDSTFNTAETGSKYPRQVNVLSHILDSMVVGCA 300
DB 241 MMSRWASHSIKIVIQIKMGAKTYPTDSTFNTAETGSKYPRQVNVLSHILDSMVVGCA 300
QY 301 MDDGGAFISWEALSLIKDLGRPKRTTLVLTAEQGGVGAFOYVQLHKVNIISVSLV 360

```

```

DB 301 MDDGGAFISWEALSLIKDLGRPKRTTLVLTAEQGGVGAFOYVQLHKVNIISVSLV 360
QY 361 MESDAGTFLPTGLQFTGSEKRAIMEEYMSLLOPLINTOVLSHEGDINFTIAGVPGA 420
DB 361 MESDAGTFLPTGLQFTGSEKRAIMEEYMSLLOPLINTOVLSHEGDINFTIAGVPGA 420
QY 421 SLDDLYKYFFHHSHGDTWTVMQKQNVAAVAVSVVADMEEMLPSS 472
DB 421 SLDDLYKYFFHHSHGDTWTVMQKQNVAAVAVSVVADMEEMLPSS 472
RESULT 13
AAB10229
ID AAB10229 standard; Protein; 472 AA.
XX AAB10229;
XX 16-NOV-2000 (first entry)
XX DE Human fetal kidney protein fragment AM282_1i.
XX Secreted protein; cytosolic; immunostimulatory; antimicrobial;
XX antiviral; immunosuppressive; antiinflammatory; vulnerary; cytokine;
XX cell proliferation; differentiation; regulator; treatment; tumor;
XX autoimmune disease; inflammatory disorder; wound; microbial infection;
XX viral disease; graft versus host reaction suppression.
XX Homo sapiens.
XX WO200037630-A1.
XX 29-JUN-2000.
XX 22-DEC-1999; 99WO-US31005.
XX 23-DEC-1998; 98US-0220876.
XX (GENY ) GENETICS INST INC.
XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR;
XX MPI: 2000-442661/38.
XX N-PSDB; AAA40493.
XX Secreted human proteing AS296-1i and AS34-1i, useful for treating
PT tumors, autoimmune diseases, inflammatory disorders, wounds, microbial
PT infections and viral diseases -
XX Disclosure; Page 199-200; 293pp; English.
XX This invention describes novel secreted human proteins (I) which have
CC cytostatic, immunostimulatory, antimicrobial, antiviral,
CC immunosuppressive, antiinflammatory and vulnerary activity and which act
CC as cytokine, cell proliferation or differentiation regulators. (I)
CC is useful for treating tumors, autoimmune diseases, inflammatory
CC disorders, wounds, microbial infections and viral diseases. (I) is also
CC useful for suppressing graft versus host reaction. AAB10226-B10288
CC represent the secreted proteins encoded by AAA40490-A40580 which are
CC described in the method of the invention.
XX
SQ Sequence 472 AA;
Query Match 99.7%; Score 2410; DB 21; Length 472;
Best Local Similarity 99.8%; Pred. No. 3e-211;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKFLIPAFGGVHLLSLCSGKAIICKNGISKRTPEEIKKEIASCGDVAKIINLVYGRQA 60
DB 1 MKFLIPAFGGVHLLSLCSGKAIICKNGISKRTPEEIKKEIASCGDVAKIINLVYGRQA 60
QY 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYQNLQDDLEKVLHPVRIIPMERGESBA 120

```

```

Db      61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESA 120
Qy      121 VMLPRRIHKIALLIGSSIGTPPEGITAEVLVWTSFDELQRRASRARKIIVVNOPTINY 180
Db      121 VMLPRRIHKIALLIGSSIGTPPEGITAEVLVWTSFDELQRRASRARKIIVVNOPTINY 180
Qy      181 SRTVQYRTQGAVEAAKVGALASLIRSVAFSISYPHTGIOEYODVPKIPACTIVEDAE 240
Db      181 SRTVQYRTQGAVEAAKVGALASLIRSVAFSISYPHTGIOEYODVPKIPACTIVEDAE 240
Qy      241 MMSRMASHGIRKIYIQLKMGAKTYPTDPSFNTVAEITSSKYPEQVVLVSGHLSMDVQGA 300
Db      241 MMSRMASHGIRKIYIQLKMGAKTYPTDPSFNTVAEITSSKYPEQVVLVSGHLSMDVQGA 300
Qy      301 MDDGGGAFISWEALSLIKDGLRPKRTLRLVMTAEBOGGVGAFOYQLHKVNISNLSLV 360
Db      301 MDDGGGAFISWEALSLIKDGLRPKRTLRLVMTAEBOGGVGAFOYQLHKVNISNLSLV 360
Qy      361 MESDAGTFLPTGLQFTGSEKARAIIMEEVMSSLOPLNTIQLVSHGEGTDINFWIQAGVPGA 420
Db      361 MESDAGTFLPTGLQFTGSEKARAIIMEEVMSSLOPLNTIQLVSHGEGTDINFWIQAGVPGA 420
Qy      421 SLDDLYKYFFPFHSHGDTMTWMDPKQNNVAAVAVVSYVADMEMLPRS 472
Db      421 SLDDLYKYFFPFHSHGDTMTWMDPKQNNVAAVAVVSYVADMEMLPRS 472

```

```

RESULT 14
AAM93559
ID      AAM93559 standard; Protein; 472 AA.
XX
AC      AAM93559;
DT      06-NOV-2001 (first entry)
XX
DE      Human polypeptide, SEQ ID NO: 3329.
XX
KW      Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS      Homo sapiens.
XX
PN      EP1130094-A2.
PD      05-SEP-2001.
XX
PF      07-JUL-2000; 2000EP-0114089.
XX
PR      08-JUL-1999; 99JP-0194486.
XX
PR      11-JAN-2000; 2000JP-0118774.
XX
PR      02-MAY-2000; 2000JP-0183765.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Ota T, Nishikawa T, Isogai T, Hayashi K, Ichii S, Kawai Y;
PI      Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR      WPI; 2001-524255/58.
XX
DR      N-PSDB; AAK94491.
XX
PT      830 Primers useful for synthesizing full length cDNA clones and their
PT      use in genetic manipulation -
XX
PS      Claim 8, SEQ ID NO 3329, 1380bp + sequence listing; English.
XX
XX      The invention relates to primers for synthesizing full length cDNA
XX      clones. 830 cDNA molecules encoding a human protein have been
XX      isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA
XX      molecules have been determined. Primers for synthesizing the full length
XX      cDNA are useful for clarifying the function of the protein encoded by
XX      the cDNA. The full length clones were obtained by construction of full
XX      length enriched cDNA libraries that were synthesised by the oligo-capping
XX      method. The primers enable the production of the full length cDNA easily

```

CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 472 AA;

Query Match 99.4%; Score 2402; DB 22; Length 472;
 Best Local Similarity 99.4%; Pred. No. 1,66-210;
 Matches 469; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 MKFLIPAFGCVHLLSLCSGKAICKNGISKRTPEEIKELIASCGDAKAIINLAVYKAQ 60
Db      1 MKFLIPAFGCVHLLSLCSGKAICKNGISKRTPEEIKELIASCGDAKAIINLAVYKAQ 60
Qy      61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESA 120
Db      61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESA 120
Qy      121 VMLPRRIHKIALLIGSSIGTPPEGITAEVLVWTSFDELQRRASRARKIIVVNOPTINY 180
Db      121 VMLPRRIHKIALLIGSSIGTPPEGITAEVLVWTSFDELQRRASRARKIIVVNOPTINY 180
Qy      181 SRTVQYRTQGAVEAAKVGALASLIRSVAFSISYPHTGIOEYODVPKIPACTIVEDAE 240
Db      181 SRTVQYRTQGAVEAAKVGALASLIRSVAFSISYPHTGIOEYODVPKIPACTIVEDAE 240
Qy      241 MMSRMASHGIRKIYIQLKMGAKTYPTDPSFNTVAEITSSKYPEQVVLVSGHLSMDVQGA 300
Db      241 MMSRMASHGIRKIYIQLKMGAKTYPTDPSFNTVAEITSSKYPEQVVLVSGHLSMDVQGA 300
Qy      301 MDDGGGAFISWEALSLIKDGLRPKRTLRLVMTAEBOGGVGAFOYQLHKVNISNLSLV 360
Db      301 MDDGGGAFISWEALSLIKDGLRPKRTLRLVMTAEBOGGVGAFOYQLHKVNISNLSLV 360
Qy      361 MESDAGTFLPTGLQFTGSEKARAIIMEEVMSSLOPLNTIQLVSHGEGTDINFWIQAGVPGA 420
Db      361 MESDAGTFLPTGLQFTGSEKARAIIMEEVMSSLOPLNTIQLVSHGEGTDINFWIQAGVPGA 420
Qy      421 SLDDLYKYFFPFHSHGDTMTWMDPKQNNVAAVAVVSYVADMEMLPRS 472
Db      421 SLDDLYKYFFPFHSHGDTMTWMDPKQNNVAAVAVVSYVADMEMLPRS 472

```

```

RESULT 15
AAM25762
ID      AAM25762 standard; Protein; 481 AA.
XX
AC      AAM25762;
DT      16-OCT-2001 (first entry)
XX
DE      Human protein sequence SEQ ID NO:1277.
XX
XX      Human; cancer; HIV infection; human immunodeficiency virus;
XX      antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
XX      antibacterial; endocrine; cardiant; central nervous system; vitruide;
XX      anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
XX      antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
XX      dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;
XX      neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX      immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
XX      antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
XX      cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
XX      genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX      thrombocytopenia; osteoporosis; severe combined immunodeficiency;
XX      allergic rhinitis; diabetes; multiple sclerosis; depression;
XX      Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX      neurological disorder.
XX
XX      Homo sapiens.
XX
XX      WO200153455-A2.

```

XX 26-JUL-2001.
PD 22-DEC-2000; 2000WO-US35017.
XX 23-DEC-1999; 99US-0471275.
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
XX (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-457603/49.
DR N-PSDB; AAH99703.
XX
XX Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS Claim 20; Page 265; 1217p; English.

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antiinflammatory; antineumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; viricide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antidiabetic; haemostatic; vulnerrary;
CC antitumor; osteopathic; dermatological; antiallergic; antiaesthetic;
CC antidiabetic; cytosolic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.

SQ Sequence 481 AA;

Query Match 98.4%; Score 2378.5; DB 22; Length 481;
Best Local Similarity 98.9%; Pred. No. 2.3e-208;
Matches 468; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 MKFLFAFFGGVHLLSLCSGKAIICNGISKRTFEIKERIASCGPAKAIINLAVYGRKQ 60
DB 5 MKFLFAFFGGVHLLSLCSGKAIICNGISKRTFEIKERIASCGPAKAIINLAVYGRKQ 68
QY 61 NRSYERLALVDTVPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESB 120
DB 65 NRSYERLALVDTVPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESB 128
QY 121 VMLERIRIKIALILGSSIGTPPEGITAEVLLVTSFDELQRRASERKGIIVYNQPIYNY 180
DB 129 VMLERIRIKIALILGSSIGTPPEGITAEVLLVTSFDELQRRASERKGIIVYNQPIYNY 188
QY 181 SRTVOYRTQGAVEAKVGLASLIRSVASFISYSPHTGIOEXYQDGVPKIPTACITVEDAE 240
DB 189 SRTVOYRTQGAVEAKVGLASLIRSVASFISYSPHTGIOEXYQDGVPKIPTACITVEDAE 248
QY 241 MMSRMAHGKIKIVIQKMGAKTYPTDSFNTVAEITGSKYRQVYLVSGHLSMDVQGA 300
DB 249 MMSRMAHGKIKIVIQKMGAKTYPTDSFNTVAEITGSKYRQVYLVSGHLSMDVQGA 308
QY 301 MDDGGGATISWALSLIKDLGLRPKRTLRVLMVTAEBQGVGAFQYQQLHKVNISNYSIV 360
DB 309 MDDGGGATISWALSLIKDLGLRPKRTLRVLMVTAEBQGVGAFQYQQLHKVNISNYSIV 368

QY 361 MESDAGTFLPTGLQFTGSEKARAIMERYMSLLOPLNITOVLSHSGGTDINFWIOAGVPGA 420
DB 369 MESDAGTFLPTGLQFTGSEKARAIMERYMSLLOPLNITOVLSHSGGTDINFWIOAGVPGA 428
QY 421 SLDDLYKYFFPHSHGDTMTWMD-PRQMNVAALVAVSVYVADMEBMLPRS 472
DB 429 SLDDLYKYFFPHSHGDTMTWMD-PRQMNVAALVAVSVYVADMEBMLPRS 481

Search completed: December 22, 2003, 11:49:03
Job time : 82.2473 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:37:50 ; Search time 38.6646 Seconds
(without alignments)
1173.984 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417
Sequence: 1 MKFLIPAFGCVHLSLSCG.....AAVAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_76:.*
2: PIR1:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759.5	31.4	493	2 T46974	leucyl aminopeptid
2	640	26.5	467	2 G87564	aminopeptidase, pr
3	182	7.5	536	2 B83278	probable aminopept
4	176.5	7.3	571	2 H87600	hypothetical prote
5	173.5	7.2	776	1 JH0570	transferrin recept
6	169	7.0	455	2 S39663	aminopeptidase hom
7	165.5	6.8	703	2 T47631	peptidase-like pro
8	165	6.8	466	2 B84132	aminopeptidase B3
9	165	6.8	750	2 A56881	prostate-specific
10	163	6.7	1483	2 T19751	hypothetical prote
11	161.5	6.7	501	2 T30154	hypothetical prote
12	158.5	6.6	501	2 G82414	aminopeptidase VCA
13	156	6.5	433	2 P84215	aminopeptidase lim
14	150.5	6.2	529	2 G82759	hypothetical prote
15	149.5	6.2	609	2 JG7819	metalloproteinase
16	149	6.1	500	2 H70629	probable AMINOPEPT
17	147.5	6.1	441	2 A84351	hypothetical prote
18	147.5	6.1	504	2 S24314	bacterial leucyl a
19	146	6.0	430	2 A87634	peptidase, M20/M25
20	143.5	5.9	622	2 A34549	transferrin recept
21	140	5.8	484	2 H90220	hypothetical prote
22	138.5	5.7	288	2 S66427	aminopeptidase (EC
23	138.5	5.7	374	2 S69639	hypothetical prote
24	137.5	5.7	537	2 A54134	aminopeptidase Y (
25	129.5	5.4	811	2 B57149	probable membrane
26	129	5.3	280	2 B5075	probable peptidase
27	126	5.2	413	2 UXHU	transferrin recept
28	125.5	5.2	413	2 S45467	aminopeptidase y h
29	122	5.0	757	2 A48592	transferrin recept

30	121.5	5.0	488	2 A87569	peptidase M20/M25/
31	117.5	4.9	1838	2 H86337	protein F5M15.26 (
32	114	4.7	474	2 E87650	peptidase, M20/M25
33	114	4.7	763	1 S29548	transferrin recept
34	113.5	4.7	794	2 T40289	hypothetical prote
35	111	4.6	411	2 A10567	allantoin amidohy
36	106.5	4.4	557	2 G72356	penicillin-binding
37	105.5	4.4	470	2 S57902	peptidase V - Lact
38	105.5	4.4	486	2 A82427	pyruvate kinase II
39	105	4.3	411	1 C64783	probable N-carbamyl
40	105	4.3	450	2 S43914	hypothetical prote
41	103.5	4.3	933	1 B48349	glycoprotein B pre
42	102	4.2	1332	2 F69732	PBSX prophage ORF
43	101.5	4.2	336	2 C83926	S-adenosylmethion
44	101.5	4.2	411	2 B71451	probable desucciny
45	101.5	4.2	470	2 AD1277	Xaa-His dipeptidase

ALIGNMENTS

RESULT 1

T46974 leucyl aminopeptidase (EC 3.4.11.-) ES-62 precursor [validated] - nematode (Acanthocheil, leucyl aminopeptidase (EC 3.4.11.-) ES-62 precursor [validated] - nematode (Acanthocheil, C/Species: Acanthocheilonema viteae C/Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 01-Feb-2002 C/Accession: T46974 R/Harnett, W.; Houston, K.M.; Tate, E.; Garate, T.; Apfel, H.; Adam, R.; Haslam, S.M.; Pe Mol. Biochem. Parasitol. 104, 11-23, 1999 A/Title: Molecular cloning and demonstration of an aminopeptidase activity in a filarial A/Reference number: Z24355, PMID:20055923; PMID:10589978 A/Accession: T46974 A/Status: preliminary; translated from GB/EMBL/DDDBJ A/Molecule type: mRNA A/Residues: 1-493 <HR> A/Cross-references: EMBL:AF071194; NID:G3386477; PIDN:AC28365.1; PID:G3386478 C/Function: A/Description: EC 3.4.11.- leucyl aminopeptidase [validated, PMID:20055923] C/Keywords: aminopeptidase

Query Match	31.4%	Score 759.5;	DB 2;	Length 493;
Best Local Similarity	36.6%	Pred. No. 2e-50;		
Matches 174;	Conservative 92;	Mismatches 166;	Indels 43;	Gaps 13;
QY	15	LSLCSGKALC-----KNGISKRTBEIKETELASGCVAKAIINLAVGAKNRSYERL	67	
DB	14	LTAVLGAALVDPDKTVAIPKNTIQTPEGEVAV-----LIQYITKGEVGLAYQWL	62	
QY	68	ALAVDTVGPRLSGSKNLEKAIQIMYQNLQODGLEKYLLEPV-RIPMERGESAVMLEPR	126	
DB	63	SKLVYDGRGHMVGSDSLKSIKIAFLBSLKNDNDKHTTEVPVPLPHVNRGNDVEMLEPR	122	
QY	127	IKHIAIIGLSSIGTPEPGITAEVLVVTSPDELQRRASEARGIIVVNPQYINVSRTVOY	186	
DB	123	NORLNVLAIG---GSEPRASATGEVTVYDDDV--KDDVRGKIIVTAQFAGYPLTKY	177	
QY	187	RTQGAVERAAKVGALASLRVASFSISYSPHTGQOEYODGPKIPTACTITYEDAEKSRMA	246	
DB	178	R-SVKLFEDQIGAVLVKSIPTFSINSPTTGAEN--TTIPAACTITEEMERLY	233	
QY	247	SHGKIYIOLKMGAKTPTDTSFNTVAIEITGSXYKPEOVVLVSGHLSMDVGGAMDGGG	306	
DB	234	RSGKIVIRMDKSHYEPINSSNLRIETGSRPSEVLLSAHVDSMDVGGALDDGAG	293	
QY	307	AFISWEALSLIKDGL---LRPRTLRVLWMTAEQGVGAFQYQLHKVNI-SNYSLV	360	
DB	294	CAVWMSALHSLIKLAERNPKEPKRTIRGIFMTSEEGYGAGKHVYITHKNDPEKEFYFV	353	
QY	361	MESDAGFLPTG---LQFTGSEKARIMEWMSLQPLNITQVL--SHBEGDINWIQ	414	
DB	354	SEDTGTGFKSTNMLAHLSFSGDKKSMRLKEITRLLSRNGIALGLNNSVQG-DVTWAK	412	
QY	415	AGVYGAALDD--LYKYFFHHSHGDTMTWMDPKQNVAAVAVSVYVADME	467	

Db 285 AMAEAERPE---GAVGKFNIAATYKJIELKTEIOKSESSNVGMIEGSDPTLKAQTIIIS 340
Qy 289 GHLDSMD-----VGGAMDGGGAFISWEALSLIKDLGPRKTRILVLTAE 337
Db 341 AHDHIGIKPKAKGCEADINNGALDNASGIATLLEVGAFGNSKVRPRKSVLVAITRAE 400
Qy 338 QGGVGAFOYQOLH---KVN1-SNYSLVME-----SDAGFLPLTGLQFTGSEKARAIME 386
Db 401 KGLVGS-DYFAVHPTVKKADIAAVNLDMPVLVLPFSDVIAF-----GSDRS----- 446
Qy 387 EVMSLLOPNTIOLSHREG-----TINWIOAGVGCASLL----- 423
Db 447 -----TIGEVKKAAGVGVIGVSDPLPEBGLFTSRSHYRFEQGVPSVFLMTGFQ 497
Qy 424 ---DDLKYFF---FHSHSGDTMTVMDEKQNNVAAVVSYVA 463
Db 498 NGGEKATFTPLKTYHNPADLKPID---YTAAARFALVNYEIA 539

RESULT 5

JH0570

transferrin receptor - chicken

C/Species: Gallus gallus (chicken)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: JH0570; S16855

R/Gerhardt, E.M.; Chan, L.N.L.; Jing, S.; Qi, M.; Trowbridge, I.S.

Gene 102, 249-254, 1991

A/Title: The cDNA sequence and primary structure of the chicken transferrin receptor.

A/Reference number: JH0570; MUID:91340160; PMID:1874449

A/Accession: JH0570

A/Molecule type: mRNA

A/Residues: 1-776 <GER>

A/Note: 581-His and 736-Gln were also found as the result of polymorphism

C/Comment: This protein mediates the endocytosis of the iron transferrin complex.

C/Superfamily: transferrin receptor

C/Keywords: glycoprotein; lipoprotein; phosphoprotein; receptor; thiolester bond; trans

F;13-22/Region: coated-pit mediated internalization signal

F;70-88/Domain: transmembrane #status predicted <TRM>

F;23/Binding site: phosphate (Ser) (covalent) #status predicted

F;70/Binding site: palmitate (Cys) (covalent) #status predicted

F;261,326,391,738/Binding site: carbonyl (Asn) (covalent) #status predicted

Query Match 7.2%; Score 173.5; DB 1; Length 776;

Best Local Similarity 22.0%; Pred. No. 4,8e-05;

Matches 87; Conservative 64; Mismatches 154; Indels 91; Gaps 19;

Qy 116 GEBSAVMLEPRHIAIKAILGLSSIGTPPEGITAEVLVTSFDELQRRASEARGLIVNQ 175
Db 215 GKEE--ILETPDAYVAVESGVSQK---YVNVYGLKKDFEIIQKVVASLNGTIVIRA 269
Qy 176 PYINYSTRV-QYRQGA-----VEAAVGLASLI---RSVASFSIYSP-----HTG 218
Db 270 GKITAEKAVANAKAGAGVLMYVDSLKYGITDILIPGHAHLGIDYDTGFPSPFNHTQ 329
Qy 219 IQEYO-DGVPIPIACTIVE-DAEWMSSM-----ASHGKIY-----IQLMG 259
Db 330 FPEPSSSLPIIAVOTISSAAARLFSKMDGDTGSEGKAIHCKVTTKESQIMVKLD 389
Qy 260 A-KTYPDTSFNTVAETIGSKYPEQVVLVSGHLDSWDVGAMDDGGGAFISWEALSLIK 318
Db 390 VNNMKRRIKINIGALIGFEPDRYVVIAGQRDSW--GPGVAKAGCTTALLLELARIYS 447
Qy 319 DL---GLRPRKTLRLVMTABEGGVGAFOYQOLAKVNISNYSLVESDAGTFLPTGLQ 374
Db 448 DIVNKGEGYKPRRSIIIFASWSAGDYAGVATGW-----LGGYSAMLAKKFTYISLDP 500
Qy 375 FTGSEKARAIMBEVMSLL-----QPLINTOVLSHGBGTDINFWIOAGVP----- 418
Db 501 VLGAHVKISAPLILYMLGSIKMGVKNPAAVSESLYKRLGPD---WKAAPPVIGLDNAA 557
Qy 419 -----GASLL-----DDLKYFFPHSHSGDTM 440

Db 558 FPLIAYSGIPVLSFGFYNKDEYR---FLDTKGDTL 590

RESULT 6

339663
aminopeptidase homolog ywad - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 07-Oct-1994 #sequence_revision 24-Feb-1995 #text_change 15-Oct-1999

C/Accession: S39663; S16427; D70050

R/Glaser, P.; Kunst, F.; Arnold, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu, I.

A/Title: Microbio1. 10, 371-384, 1993

A/Reference number: S39663; MUID:95020537; PMID:7934828

A/Accession: S39663

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-455 <GLA>

A/Cross-references: EMBL:X73124; NID:9413923; PIDN:CA51564.1; PID:9413932

R/Glaser, P.; Kunst, F.; Debarbouille, M.; Vertes, A.; Danchin, A.; Dedonder, R.

DNA Seq. 1, 251-261, 1991

A/Title: A gene encoding a tyrosine tRNA synthetase is located near sacs in Bacillus sub

A/Reference number: S16421; MUID:92216127; PMID:11806041

A/Accession: S16427

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-68 <GL2>

A/Cross-references: EMBL:X52480; NID:940236; PIDN:CA56725.1; PID:940243

R/Kunst, F.; Ogatawa, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berto

C./Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

iech, U.; Harwood, C.R.; Henaux, A.; Hilbert, H.; Holleppel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lathiois,

A/Authors: Lander, U.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

V, M.; Ogawa, K.; Ogatawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadele, Y.; Sato, T.; Scanlon,

A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowka, A.; Seror

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetz, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winder, P.; Wipet, A.; Yamamoto, H.; Yoshikawa, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: D70050

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-455 <GBN>

A/Cross-references: GB:299123; GB:AL009126; NID:92636240; PIDN:CAB15873.1; PID:el186346;

A/Experimental source: strain 168

A/Genetic:

A/Gene: ywad

Query Match 7.0%; Score 169; DB 2; Length 455;

Best Local Similarity 23.2%; Pred. No. 4,7e-05;

Matches 100; Conservative 83; Mismatches 174; Indels 74; Gaps 22;

Qy 58 KAQRSTBRLLALVDYTPRISGSKNLEKAIQIYQULQODGLEKHLPRIRHMERGE 117
Db 44 KAKH-AVSTISQLEAIGPRAGTAAKSKMLIASSMRKIKLD-VVQQRNIDRLREGT 101
Qy 118 ESAVMLEPRHIAIKAILGLSSIGTPPEGITAEVLVTSFDELQRRASEARGLIVNQ 177
Db 102 LSSAGRD---LILQASGAGAPTEBGLTAP-LYNNAGLGVOKOPTADAKKILISGCD 155
Qy 178 INVSRTVQYRTQGAVEAKAGALASLI-RSVAASFISYSPHTGQYODGVPIPIACTIV 236
Db 156 LTY-----YEAKAAEAA--GAKAVIILYNNKESLVPTPNLSGKV-----GLVVGKK 203
Qy 237 EDAMSRMASHGKIYIYQLKMGAKTYPTDTSFNTVAETIGSKYPE-----QVLYVSGHL 291
Db 204 EDGALTOQKEATIKL-----KAFNTQTSQN-----IIGIKPKKNIKHPDIYVYAHY 251

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-750 <ISR>
 A:Cross-references: GB:999487; NID:g190663; PIDN:AAA60209.1; PID:g190664
 A:Experimental source: prostatic carcinoma cell line LNCaP
 A>Note: sequence extracted from NCBI backbone (NCBIN:121724, NCBIPI:121725)
 C:Superfamily: transferin receptor
 C:Keywords: surface antigen; transmembrane protein

Query Match 6.8%; Score 165; DB 2; Length 750;
 Best Local Similarity 19.2%; Pred. No. 0.0002;
 Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

```

Qy 3 FLTFAFGVHLLSLCSGKAIKNGISKRTFEIKEEIASCGDVAKAIINLAVYGAQNR 62
D 32 FLLGLFLEGMFKISNENATNTPKHNKAFDELKAE-----NKKFLVNF----- 77
Qy 63 SYERLALLVDVTPGRLSGSK--NLEKAIQIMYQNLQODGLEKVLBPV-----RIPHW 113
D 78 -----TQIDPLAGTEQNQLAKQIQGQKKEFGDVELAHVDLLSYPKTHPNY 127
Qy 114 -----ERGEE--SAVMEPRIRHKAILGLSSIGTPPGITAEVLVVTSPELQRRSEA 166
D 128 ISIINEGNEIFENSLFEP-----PPGYENVSDIVPPGSAFSPQCM-P 170
Qy 167 RGIIVVNOPPIYNSRT-----VOYRTOGAVEAKVG-----ALASLRSV 207
D 171 EGDLY-----YVYAKTEDPFLERDMKINSGLIVARVGRKVKVKAQLAGAKV 225
Qy 208 ASFS-----IYSPHTGIOEYOD-----ALASLRSV 224
D 226 ILVSDPADVYFAP--GVKSPYDGMNLPGGVQGRNINLNGADPLTPGYPANEAAYRGI 283
Qy 225 -----GVKRIPIACTVEDAEKMSMAHSIGIKIVQLKMGATYD----- 265
D 284 AEAVGLSPVHPHIGYDDAKLE-----KMGSAAPPDSSWRGSLKVPYVNGP 331
Qy 266 -----TDSFNTVAITGSKYPEOVVLVSGHLDSDMDVGQAMDG 304
D 332 GFTGNSTQKVKMHIHNEVTRYVNGTIRGAVEPRVYIILGHRDSWF--GGIDPQ 389
Qy 305 GGAFISWE--ALSLIDQLGPRKTLVLTAEBOGVGAFOYQLHKVNIYSYLV 361
D 390 SGAIVVHIVRSFGLKKEGWRPRRTILFASMDAEEFLGSTWAE-----ENSRLLQ 443
Qy 362 ESDAGTFLPTGLQFTGSEKARIMEEVM--SLQPLNTIYVLSHSGDIDINFIQAVPGA 420
D 444 ERGV-AYINADSIEGNYTLRVDCITPLMYSLVH--NLTKELKSPD-----EGEGRK 491
Qy 421 SLDDLK 428
D 492 SLYESWK 499

```

RESULT 10
 T19751
 hypothetical protein C35C5.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T19751
 R:White, S.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: Z19173
 A:Accession: T19751
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1483 <WLL>
 A:Cross-references: EMBL:Z78417; PIDN:CAB01688.1; GSPDB:GN00028; CESP:C35C5.2
 A:Experimental source: clone C35C5
 C:Genetic8:
 A:Gene: CESP:C35C5.2
 A:Map position: X
 A:introns: 61/1; 203/3; 248/1; 293/3; 331/1; 437/1; 485/2; 523/2; 556/1; 595/3; 620/2; 6

Query Match 6.7%; Score 163; DB 2; Length 1483;
 Best Local Similarity 19.5%; Pred. No. 0.0003;
 Matches 108; Conservative 82; Mismatches 181; Indels 184; Gaps 24;

```

Qy 25 KNGISKRTFEIEE--IASCQVAKAIINLA--YVGAQNRSYE-----RLALLVD 72
D 7 KHTFLQKVFQKSRKNGMKASGVVLAIVSTALTIILSNAIHOSYKSNKFLPKLSIATV 66
Qy 73 -----TVGPRSSGSKNLEKAIQIMYQNLQODGLEKVLBPV-----P 111
D 67 IKOLINVDNIRSNLHSLTKKPHVAGTENNIRVEMIRDOITQGLENVHNEYVMLSYP 126
Qy 112 HMERGESAVMLPRIRHKAILGLSSIGTPPGITAEVLVVTSFD--ELQRRSEARG 168
D 127 NW-----TTPNI--IEILKENDDIVSTTGRSVVKEQNDPLAEIQWLAISAAG 175
Qy 169 KI---VVY---NOPPIYNSRTV-----OYRTOGAVEAKVAGLASLI-- 204
D 176 TVEGDIYVNNANPSDIEYLESIGIDLKDKIFLARYSSNRYGNIQAQVAKKACLVYS 235
Qy 205 --RSVASF-----SIY-----SPTGT-----OEYQ 223
D 236 DPRQVASLGTGPNETYGNTDKMPSNTVQMGSVYIGLDPETPAFSGIDLFKEKTEGDDL 295
Qy 224 D--GVKRIPIACTVEDAEKMSMAHSIGIKIVQLKMGATYD-----SRMASHG 249
D 296 DEKRIPIPLPITVYATQILFENMKGDVADNAPQGLNATYRGRPLINNQLRVYVHA 355
Qy 250 -----IKIVIQLMGAKTYPDTSFNTVAITGSKYPEOVVLVSGHLDSDMDVQGA 300
D 356 ENEERSVIVVELLSRLINFRKIQ-----NIMGVIKSGSPDRKVFVLSNHYDAWT--YGA 408
Qy 301 MDGCGAFIWEALSLIK---DLGLRPRKTLVLTAEBOGVGAFOYQLHKVNIYS 356
D 409 VDPNSGTSTLLEVSRAIKQYQNTGMPARSILFAMHDAEYGLIGSTFAEERYQLMR 468
Qy 357 YSL-VNESDAGTFLPTGLQFTGSEKARIMEEVM--SLQPLNTIYVLSHSGDIDINFIQ 415
D 469 RAVAVIMMDLIGANQTLIGISNFTVANVLASAANVEQ--NPTBM----- 512
Qy 416 GVPGASLDDLK 430
D 513 -EGRKTLVDSWKY 526

```

RESULT 11
 T30154
 hypothetical protein R57.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T30154
 R:Favella, T.; Rifkin, L.; Chiapelli, B.
 submitted to the EMBL Data Library, February 1997
 A:Description: The sequence of C. elegans cosmid R57.
 A:Reference number: Z20745
 A:Accession: T30154
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-751 <FAV>
 A:Cross-references: EMBL:U88179; PIDN:AB52660.1; GSPDB:GN00028; CESP:R57.1
 A:Experimental source: strain Bristol N2; clone R57
 C:Genetic8:
 A:Gene: CESP:R57.1
 A:Map position: X
 A:introns: 40/1; 145/1; 230/1; 275/3; 313/1; 344/2; 405/1; 453/2; 576/2; 671/2; 713/3

Query Match 6.7%; Score 161.5; DB 2; Length 751;
 Best Local Similarity 18.8%; Pred. No. 0.00038;
 Matches 118; Conservative 108; Mismatches 188; Indels 213; Gaps 33;
 4 LIPAFPGVHLLSLCSGKAIKNGI--SKRTFEIKEEIASCGDVAKA---INLAVYGA 58

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:34:20 ; Search time 29.9104 Seconds
(without alignments)
742.104 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417
Sequence: 1 MKELIFAFGCVHLLSLCSG.....AVNAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	188	7.8	740	1	NID2_HUMAN
2	176	7.3	752	1	FOH1_RAT
3	175.5	7.3	776	1	TPR1_CHICK
4	174	7.2	752	1	FOH1_MOUSE
5	172.5	7.1	705	1	GCPT2_BACSU
6	169	7.0	455	1	YRAD_BACSU
7	166	6.9	751	1	FOH1_PIG
8	165	6.8	750	1	FOH1_HUMAN
9	161.5	6.7	751	1	GCPT2_MOUSE
10	153.5	6.4	769	1	TRP1_FELCA
11	147.5	6.1	504	1	AMPX_VIBR
12	143.5	5.9	622	1	TRP1_RAT
13	143.5	5.9	740	1	NID2_HUMAN
14	142	5.9	798	1	TRP2_MOUSE
15	138.5	5.7	284	1	APX_STRGR
16	137.5	5.7	537	1	APB3_YEAST
17	132	5.5	801	1	TRP2_HUMAN
18	130.5	5.4	770	1	TRP1_CANFA
19	129.5	5.4	811	1	YV96_YEAST
20	126	5.2	760	1	TRP1_HUMAN
21	125.5	5.2	413	1	YBS4_YEAST
22	123.5	5.1	745	1	NID1_RAT
23	122	5.0	757	1	TRP1_CRICR
24	114	4.7	763	1	TRP1_MOUSE
25	113.5	4.7	794	1	YV99_SCHPO
26	110.5	4.6	456	1	SR54_THEAC
27	109.5	4.5	541	1	PTCD_MOUSE
28	105.5	4.4	470	1	PEPV_LACDL
29	105	4.3	411	1	ALIC_ECOLI
30	102	4.2	1332	1	XKDO_BACSU
31	101	4.2	490	1	RBL_ODOSI
32	100.5	4.2	933	1	VGLB_HSVAI
33	99	4.1	416	1	PGK_GLOMO

34	99	4.1	843	1	NUOG_STRCO	Q9xar0 streptomyc
35	98.5	4.1	809	1	DCML_OLICA	P19919 oligotropha
36	98.5	4.1	1396	1	VCAP_VZVD	P09245 varicella-z
37	98.5	4.1	1485	1	TOP2_SCHPO	P08096 schizosacch
38	97.5	4.0	145	1	CBPG_PSSS6	P06621 pseudomonas
39	97	4.0	886	1	YIH3_YEAST	P40511 saccharomyc
40	96.5	4.0	811	1	HSP7_YEAST	P33416 saccharomyc
41	96	4.0	385	1	SUCC_BACSU	P80886 bacillus su
42	96	4.0	693	1	GUAA_HUMAN	P49915 homo sapien
43	95.5	4.0	340	1	SRR_HUMAN	O99zt4 homo sapien
44	95.5	4.0	541	1	PTCD_RAT	O88618 r formimido
45	95	3.9	368	1	FTZ1_ARCFU	O29715 archaeoglob

ALIGNMENTS

RESULT 1
NID2_HUMAN STANDARD; PRT; 740 AA.
AC Q9Y3Q0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE N-acetylated-alpha-linked acidic dipeptidase II (EC 3.4.17.21)
DE (NALADase II).
GN NALAD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI Taxid=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung carcinoma;
RX MEDLINE=9185063; Pubmed=10085079;
RA Pangalos M.N., Neels J.-M., Somers M., Verhaesselt P., Bekkers M.,
RA van der Helm L., Fraiponte E., Ashton D., Gordon R.D.;
RT "Isolation and expression of novel human glutamate carboxypeptidases
RT with N-acetylated alpha-linked acidic dipeptidase and dipeptidyl
RT peptidase IV activity.";
RL J. Biol. Chem. 274:8470-8483(1999).
CC -1- FUNCTION: Has N-acetylated-alpha-linked-acidic dipeptidase
CC (NALADase) activity. Also exhibits a dipeptidyl-peptidase IV type
CC activity. In vitro, cleaves Gly-Pro-AMC.
CC -1- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal
CC glutamyl residue, typically from Ac-Aap-Glu or pteroyl-gamma.
CC -1- COFACTOR: Zinc; Binds two ions per subunit. Required for NALADase
CC activity.
CC -1- ENZYME REGULATION: Inhibited by quisqualate.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Highest expression in the testis. Also found
CC in ovary and spleen. Weak expression in prostate, heart and
CC placenta. In brain, expressed in striatum, parietal cortex and
CC ventral striatum with lower levels in hippocampus, brain stem,
CC putamen and superior colliculus.
CC -1- DOMAIN: The NALADase activity is found in the central region, the
CC dipeptidyl peptidase IV type activity in the C-terminal.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi-sib.ch).
CC
CC EMBL: AJ012370; CAB39967.1; -
CC MEROPS: M28.012; -
CC Genew: HGNC:14526; NALAD2.
CC GO: GO:0016021; C:integral to membrane; NAS.
CC GO: GO:0004180; F:carboxypeptidase activity; NAS.
CC GO: GO:0016805; F:dipeptidase activity; NAS.

DR GO: GO:0004274; F: dipeptidyl-peptidase IV activity; NMS.
 DR GO: GO:0006508; P: proteolysis and peptidolysis; NMS.
 DR InterPro: IPR00137; PA.
 DR Pfam: PF02225; PA; 1.
 DR Pfam: PF04389; peptidase_M28; 1.
 DR Pfam: PF04253; TFR dimer; 1.
 KW Hydroxylase; Carboxypeptidase; Metalloprotease; Zinc; Aminopeptidase;
 KW Dipeptidase; Serine protease; Transmembrane; Glycoprotein;
 KW Signal-anchor; Multifunctional enzyme.
 FT DOMAIN 1 7
 FT TRANSMEM 8 31
 FT
 FT DOMAIN 32 740
 FT DOMAIN 264 577
 FT ACT_SITE 414 414
 FT ACT_SITE 618 618
 FT ACT_SITE 656 656
 FT ACT_SITE 679 679
 FT METAL 367 367
 FT METAL 377 377
 FT METAL 443 443
 FT METAL 443 443
 FT METAL 543 543
 FT CARBOHYD 111 111
 FT CARBOHYD 143 143
 FT CARBOHYD 185 185
 FT CARBOHYD 314 314
 FT CARBOHYD 449 449
 FT CARBOHYD 603 603
 FT CARBOHYD 628 628
 SQ SEQUENCE 740 AA; 83591 MW; 040624D691ECF879 CRC64;

Best Match 7.8%; Score 188; DB 1; Length 740;
 Query Local Similarity 20.0%; Pred. No. 6,9e-06;
 Matches 112; Conservative 85; Mismatches 186; Indels 176; Gaps 27;
 58 KAON-----RSYERLALVDTVGPRLSGSKN---LEKAIQIYQNIQODGLEKHLER-V 108
 56 KAENKSFLESTFKL-----PHLAGTEQNFLLAKKIQTOMKRGFSDASLVVYDVL 107
 109 RHPWEEBSAVMLEPRIKIALIGLSSIGTPEGITAEVLVVTSDDELORRSEARG 168
 108 SYPNETNANYISIVDE---HETEIFKT-SYLEPPDGEYNTNIVPYNFAFGQM-DEG 162
 169 KIVVNOYINYSRVVQY-----RPGAVEAKKAL-----ASLRVYASPSI 212
 163 DLV-----YVNYARTDEPFKLEREMGINCTGKIIVARYGKIFRGKVKVNMALAGIIL 217
 213 YS-----PHTGIQ-----EYOD 224
 218 YSDPADYAPRYQYPRPKGMNLPETIAQKQNVNLNAGADPLTRGYPAKEYTRFLDVEGCV 277
 225 GVPKIPTACTIVEDAEMMSRMASHGIKIVIQKMAKT-----YPTDSF----- 269
 278 GIPRIPIVPIGVNDAEILRLRYLG-GIAPDKSMKALNVYSIGFGFSGDSFRKRVNHV 336
 270 -----NTVAITTSKGYEYQVVLVSGHDSMDVQGANDDGAGFISME---ALSL 316
 337 YNINKITRIYVNVGTRGSVEPDRVYLIGGHDVWF-GALDPTSGVAVLQELIARSPGK 394
 317 IKDLGRLPKRTLRVLMTAEBOGUGAFQYQVLHKVNISNYSVLWESDAGTFLPTGLPT 376
 395 LMSKMRRTIRTIIPASWDAEERGLLGSTEMAE-----ENVKILDERST-AIINDSSIE 447
 377 GSEKARA-----IMEEYVSLQPL-----NITQVLSHGE 405
 448 GNVTARVQVCTPLLYLVVLTKEIPSPDDGFESKSLYESMLEKDESPENKNLPRIKXLS 507
 406 GTDIFWIO-AGVPGA-----SLDDLYKYFFHHSHDQDTTWV-----DP---KOMVYA 451
 508 GSDFAFYFQRLGIASGARARYTKKTKDKYSYVYHTIYETELVEKFPYDPFKKOLSA 567
 452 AAVMAVSVVADMEEMLP 470

DB 568 QLRGALV-YELVD-SKILP 584
 RESULT 2
 FOHL RAT STANDARD; PRT; 752 AA.
 ID FOHL RAT
 AC P70627;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 42, Last annotation update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glutamate carboxypeptidase II (EC 3.4.17.21) (Membrane glutamate
 DE carboxypeptidase) (MGCP) (N-acetylated-alpha-linked acidic dipeptidase
 DE I) (NAAADase I) (Pteroylpolymy-gamma-glutamate carboxypeptidase)
 DE (Folyl)polymy-gamma-glutamate carboxypeptidase (FGCP) (Folate hydrolase
 DE 1) (Folate-specific membrane antigen homolog).
 GN FOHLI OR NAAAD1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RX MEDLINE=98041505; PubMed=9375657;
 RA Bzdega T., Turi T., Wroblewska B., She D., Chung H.S., Kim H.,
 RA Neale J.H.,
 RT "Molecular cloning of a peptidase against N-acetylaspartylglutamate
 RT from a rat hippocampal cDNA library."
 RT J. Neurochem. 69:2270-2277 (1997).
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Hippocampus;
 RX MEDLINE=9816924; PubMed=9501243;
 RA Luchli-Carter R., Berger U.V., Barczak A.K., Emma M., Coyte J.T.;
 RT "Isolation and expression of a rat brain cDNA encoding glutamate
 RT carboxypeptidase II."
 RT Proc. Natl. Acad. Sci. U.S.A. 95:3215-3220 (1998).
 RN [3]
 RP SEQUENCE OF 284-752 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96149377; PubMed=8570628;
 RA Carter R.E., Feldman A.R., Coyte J.T.;
 RT "Prostate-specific membrane antigen is a hydrolase with substrate and
 RT pharmacologic characteristics of a neuropeptidase."
 RT Proc. Natl. Acad. Sci. U.S.A. 93:749-753 (1996).
 RN [4]
 RP SEQUENCE OF 284-752 FROM N.A.
 RC TISSUE=Brain;
 RA Luchli-Carter R., Berger U.V., Barczak A.K., Emma M., Coyte J.T.;
 RT Submitted (DBS-1997) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP ALTERNATIVE SPLICING.
 RA Bzdega T., She D., Turi T., Wroblewska B., Neale J.H.;
 RT "Molecular cloning of alternatively spliced variants of the peptidase
 RT against N-acetylaspartylglutamate (NAAAG) from human and rat nervous
 RT systems."
 RL Abstr. Soc. Neurosci. 24:579-579 (1998).
 CC -1- FUNCTION: Has both folate hydrolase and N-acetylated-alpha-linked-
 CC acidic dipeptidase (NAAADase) activity. Has a preference for tri-
 CC alpha-glutamate peptides (By similarity). In the intestine,
 CC required for the uptake of folate. In the brain, modulates
 CC excitatory neurotransmission through the hydrolysis of the
 CC neuropeptide, N-acetylaspartylglutamate (NAAAG), thereby releasing
 CC glutamate.
 CC -1- FUNCTION: Also exhibits a dipeptidyl-peptidase IV type activity
 CC (By similarity). In vitro, cleaves Gly-Pro-AMC (By similarity).
 CC -1- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal
 CC glutamyl residue, typically from Ac-Asp-Glu or pteroyl-gamma-
 CC -1- COFACTOR: Zinc; Binds two ions per subunit. Required for NAAADase
 CC activity.
 CC -1- ENZYME REGULATION: The NAAADase activity is inhibited by beta-
 CC NAAAG, quisqualic acid and 2-(phosphonomethyl)glutaric acid (PMG).

```
CC CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Plasma membrane  
CC CC (by similarity).  
CC CC -I- ALTERNATIVE PRODUCTS:  
CC CC Event=Alternative splicing; Named isoforms=3;  
CC CC Comment=Experimental confirmation may be lacking for some  
CC CC isoforms;  
CC CC Name=1;  
CC CC IsoId=P70627-1; Sequence=Displayed;  
CC CC Name=2; Synonyms=Short form;  
CC CC IsoId=P70627-2; Sequence=Not described;  
CC CC Note=Probably inactive;  
CC CC Name=3; Synonyms=Long form;  
CC CC IsoId=P70627-3; Sequence=Not described;  
CC CC -I- TISSUE SPECIFICITY: Widely expressed throughout brain regions with  
CC CC highest levels in the hippocampus, dentate gyrus, piriform cortex,  
CC CC choroid plexus of ventricles, pineal gland, anterior lobe of the  
CC CC pituitary gland and supraprostatic nucleus. High levels also found in  
CC CC the cerebral cortex, substantia nigra, pontine nucleus and the  
CC CC granule cell layer of cerebellum. Highly expressed in astrocytes  
CC CC and non-myelinating Schwann cells. Also expressed in kidney.  
CC CC localizing to the proximal brush border of the renal tube.  
CC CC -I- DOMAIN: The NAAALDase activity is found in the central region, the  
CC CC dipeptidyl peptidase IV type activity in the C-terminal.  
CC CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.  
-----  
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC CC the European Bioinformatics Institute. There are no restrictions on its  
CC CC use by non-profit institutions as long as its content is in no way  
CC CC modified and this statement is not removed. Usage by and for commercial  
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC CC or send an email to license@isb-sib.ch).  
-----  
DR DR EMBL; U75973; AAC53423.1; -.  
DR DR EMBL; AF040256; AACQ0067.1; -.  
DR DR EMBL; AF039707; AAB96759.1; -.  
DR DR MEROPS: M28_010; -.  
DR DR InterPro; IPR003137; PA.  
DR DR Pfam; PF02225; PA; 1.  
DR DR Pfam; PF04389; Peptidase_M28; 1.  
DR DR Pfam; PFO4253; TFR_dimer; 1.  
KW KW Hydrolyse: Carboxypeptidase; Metalloprotease; Zinc; Aminopeptidase;  
KW Dipeptidase; Serine protease; Transmembrane; Glycoprotein;  
KW Signal-anchor; Multifunctional enzyme; Alternative splicing.  
FT FT DOMAIN 1 19  
FT FT TRANSMEM 20 44  
FT FT (POTENTIAL).  
FT FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT FT (POTENTIAL).  
FT FT EXTRACELLULAR (POTENTIAL).  
FT FT NAAALDASE.  
FT FT NUCLEOPHILE (NAAALDASE) (BY SIMILARITY).  
FT FT CHARGE RELAY SYSTEM (POTENTIAL).  
FT FT ACT_SITE 426 436  
FT FT ACT_SITE 630 630  
FT FT ACT_SITE 668 668  
FT FT ACT_SITE 691 691  
FT FT METAL 379 379  
FT FT METAL 389 389  
FT FT METAL 427 427  
FT FT METAL 455 455  
FT FT METAL 555 555  
FT FT METAL 555 555  
FT FT ZINC 1 (BY SIMILARITY).  
FT FT ZINC 1 AND 2 (BY SIMILARITY).  
FT FT ZINC 2 (BY SIMILARITY).  
FT FT ZINC 1 (BY SIMILARITY).  
FT FT ZINC 1 (BY SIMILARITY).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT N-LINKED (GLCNAC . . ) (POTENTIAL).  
FT FT CARBOHYD 615 615  
FT FT CARBOHYD 640 640  
SQ SQ SEQUENCE 752 AA; 84539 MW; 5C0915A3B9C71E41 CRC64;
```

```

OY 3 FLFAEFGGHLILSCSGAKICNKSIRTEPEIKEEELASCGDAKKIIMLAIVGKAQNR 62
Db 34 FIIIGFLPGFPIKPSNDSTSSVSPGMKKAFLDELKAE-----NIKKFLYNF----- 79
OY 63 SYERLALLVDVGPRLSGSK-NLEKAIQIMYQNLQODGLEKYLHPVRI-----PH 112
Db 80 -----TRPHLAGTQHNFEILAKQI-HAQMKEFGLDELVELSDVDLLSYDNKTHPN 128
OY 113 W-----ERGBESAWMLEPIKIKALIGLSSISGTRP-EGITAEVLVYTSDELQRASE 165
Db 129 YISIIINEDGNE-----IFKTSIAEL-----SPGYESIISDVVPYSARSP-----QGT 171
OY 166 ARKIVVYNOQPIYNSRT-----VOYRQGVAEAKGAL-----ASLIRSVAS 209
Db 172 PEGDVLV-----YVNARTEDPFKLERWKINKINSGKIYIARIGQVPRGNKVNAGLAAGK 228
OY 210 FSIIYSPHT-----GIGQYOD----- 224
Db 227 IILYSDPADYFVVGKSPYDPGNLPGGGVQGRQNVNLNLAGADPLTPGPANVAYRHEFT 286
OY 225 ---GVKPIPLACTYVEDAE-MMSRAASH-----GIKIVI-----QLKM 258
Db 287 EAVGLPSIIHPHGYDDAQKLEHNHGSAPDPSNKGGLKVPYVNGPGFAGNFSKOKVKL 346
OY 259 GAKTYPD-TDSFTVAIEITGSKYPCQVVLVSGHLDSMPVGGAMDGSGAFISWE--AL 314
Db 347 HHSYKNTKRIYINVICTLKGAVBPRRYVILGHRDAPW-F--GGIDPQSGAAVHETRTF 404
OY 315 SLIKDGLRPKRTLRILVMTAEBOGVGAFOYYOLHKVYNISNYSLVMSBAGTFLPTGLQ 374
Db 405 GTLKKKGMPRRRTILPASWDAREFGILGSTWAEFH-----SRLLQGRVAVINADSS 457
OY 375 FTQSEKARALMEWSMLQPLNITQVLSHGETDINFIQGVGASILD 424
Db 458 IEGNVTLRVDCTPLMHSL-VNVLTKELSPD-----EGFEKSLYD 497

RESULT 3
TFR1_CHICK
ID_TFR1_CHICK STANDARD; PRT; 776 AA.
AC OQ09097;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transferrin receptor protein 1 (TfR1) (Tr) (TfR) (TfRf).
GN TFR1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Erythroblast, and lymphoma;
RX MEDLINE=91340160; PubMed=1874449;
RA Gerhardt E.M., Chan L.-N.L., Jung S., Qi M., Trowbridge I.S.;
RT "The cDNA sequence and primary structure of the chicken transferrin
RT receptor. ";
RL Gene 1021249-254(1991).
CC -I FUNCTION: Cellular uptake of iron occurs via receptor-mediated
CC endocytosis of ligand-occupied transferrin receptor into
CC specialized endosomes. Endosomal acidification leads to iron
CC release. The apotransferrin-receptor complex is then recycled to
CC the cell surface with a return to neutral pH and the concomitant
CC loss of affinity of apotransferrin for its receptor. Transferrin
CC receptor is necessary for development of erythrocytes and the
CC nervous system (By similarity).
CC -I SUBUNIT: Homodimer; disulfide-linked. Binds one transferrin
CC molecule per polypeptide chain (By similarity).
CC -I SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -I PTM: N- and O-glycosylated, phosphorylated and palmitoylated (By
CC similarity).

```


-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
 -1- CAUTION: A frameshift has been introduced after position 760 to produce a C-terminal sequence homologous to other species transfeitin receptor. The correct sequence is shown in fig.1 of Ref.1.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X55348; CA339035.1; ALT_FRAME.

DR HSBP; P02786; ICM8.

DR InterPro; IPR003137; PA.

DR Pfam; PF04389; Peptidase M28; 1.

DR Pfam; PF04253; TFR dimer; 1.

KM Transmembrane; Glycoprotein; Receptor; Lipoprotein; Palmitate;
 KM Signal-anchor; Endocytosis; Phosphorylation; Polymorphism.

FT TRANSMEM 1 70
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).

FT DOMAIN 92 776
 FT LIGAND-BINDING (POTENTIAL).
 FT ENDOTOXIN SIGNAL.

FT SITE 19 22
 FT CELL ATTACHMENT SITE (POTENTIAL).

FT MOD RES 23 23
 FT PALMITATE (BY SIMILARITY).

FT DIST RES 101 101
 FT PHOSPHORYLATION (BY SIMILARITY).

FT CARBOHYD 261 261
 FT INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 326 326
 FT N-LINKED (GLCNAC. . .) (BY SIMILARITY).

FT CARBOHYD 331 391
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 738 738
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARIANT 581 581
 FT R -> H (IN BUREAL LYPHOMA).

FT VARIANT 736 736
 FT K -> Q (IN BUREAL LYPHOMA).

FT SEQUENCE 776 AA; 85658 MW; DB293BAC41ED0623 CRC64;

Query Match 7.3%; Score 175.5; DB 1; Length 776;
 Best Local Similarity 21.8%; Pred. No. 5.9e-05;
 Matches 83; Conservative 58; Mismatches 142; Indels 97; Gaps 18;

CC 136 GSSITPEPGITAEVLVNTSPDELGRASERAGKIVNPNINSRYQRTQCAVNA 195
 CC 233 GVSQSKP---VYVNYGLKKDFEIIQKVASLNGTIVVRAKGI---TLAEKVANAKKAG 285
 CC 196 KVGALA---SLIRSVASFSI-----YSP-----HTGIGEOYQ-DGVPKIPATC 233
 CC 286 AAGLVMTYDLSLMTGTTDLIPFGHAHLGDPRTTGFSFNHTTQPPVBSGLPHIANQT 345
 CC 234 ITVE-DAEMMSRM-----ASHGIKIV-----IOLMGA-KTYPPTDSFNTVAE 274
 CC 346 ISSAAALRFSKMDPTCESEGWKGAHSCVTKTQESQIMVYLDVNMKBRKILNIFGA 405
 CC 275 ITGSKTPQOVVLVSHLDSWDVGQAMDDGCAFISSWALSILKYL-----GLRPKTRRL 330
 CC 406 IQGFEPFRYVVIQQRDSW--GPEVAKAGTITALLBLARVTSIVNKEGKPPRSITIF 463
 CC 331 VLMTAEQGVGAFQYVQLHKVNIYNYSLVMSDAGTFLPTQLGTGSEKARAMEEYWS 390
 CC 464 ASMSGADVGAVATW-----LEGYSAMLAHAKFTYISLDAPVLGASHVYKISAPPLY 516
 CC 391 LL-----OPTNITOVLSHGBSTDNFMIOAGVP-----GASLL----- 423
 CC 517 MLTGISIMGVKNPAVAVSSESLVNRLLPDD--WTKAVVPLGLDNAAAPPLAVSGIPVLSGPF 573
 CC 424 ---DLKYKFFPHSHSGDTM 440
 CC 574 YNKDEBYR---FLDTKGDTL 590

RESULT 4

FOH1 MOUSE

ID FOH1 MOUSE STANDARD; PRT; 752 AA.

AC O35409; Q9DCC2; 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glutamate carboxypeptidase II (BC 3.4.17.21) (Membrane glutamate

DE carboxypeptidase) (MGCP) (N-acetylated-alpha-linked acidic dipeptidase

DE I) (NALADase I) (Pteroylpolypoly-gamma-glutamate carboxypeptidase)

DE 1) (Prostate-specific membrane antigen homology) (PGCP) (Folate hydrolase

GN FOH1 OR MOPSM OR NALAD1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

ON [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NIH Swiss; TISSUE=Brain;

RX MEDLINE=21077532; PubMed=11210180;

RT "Cloning, expression, genomic localization, and enzymatic activities

RT of the mouse homolog of prostate-specific membrane antigen/NALADase/
 RT folate hydrolase.";

RT Mamm. Genome 12:117-123 (2001).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Kidney;

RX MEDLINE=21085660; PubMed=11217851;

RA Arakawa T., Hara A., Fukunishi Y., Yoshino M., Itoh M., Ienhi Y.,
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Kono H., Baldairelli R., Baren G.,
 RA Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshew-Borle A., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,
 RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

CC -1- FUNCTION: Has both folate hydrolase and N-acetylated-alpha-linked-
 CC acidic dipeptidase (NALADase) activity. Has a preference for tri-
 CC alpha-glutamate peptides (By similarity). In the intestine,
 CC required for the uptake of folate. In the brain, modulates
 CC excitatory neurotransmission through the hydrolysis of the
 CC neuropeptide, N-acetylaspartylglutamate (NAAG), thereby releasing
 CC glutamate.CC -1- FUNCTION: Also exhibits a dipeptidyl-peptidase IV type activity
 CC (By similarity). In vitro, cleaves Gly-Pro-AMC (By similarity).CC -1- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal
 CC dipeptidyl residue, typically from Ac-Asp-Glu or pteroyl-gamma-CC -1- COFACTOR: zinc; Binds two ions per subunit. Required for NALADase
 CC activity.CC -1- ENZYME REGULATION: The NALADase and folate hydrolase activities
 CC are inhibited by quiescent acid.CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Plasma membrane
 CC (By similarity).CC -1- TISSUE SPECIFICITY: Expressed predominantly in the hippocampal
 CC region of the brain and in kidney. Lower levels in the ovary,
 CC testis and mandibular gland.CC -1- DOMAIN: The NALADase activity is found in the central region, the
 CC dipeptidyl peptidase IV type activity in the C-terminal.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.

-1- CAUTION: There are amino acid differences between the sequence shown in fig.1 (Ref.1) and the sequence deposited in the database (AF026380). The sequence from fig.1 shows only 3 conflicts between Ref.1 and Ref.2. These are at AA positions 141, 240 and 287.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as the content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF026380; AAB81971.1; -
 DR EMBL; AK002920; BAB22457.1; -
 DR MEROPS; M28.010; -
 DR MGD; MGI.1858193; Fc1h1.
 DR GO; GO:0005887; C:integral to plasma membrane; IDA.
 DR GO; GO:0008464; F:gamma-glutamyl hydrolase activity; IDA.
 DR GO; GO:0004237; F:membrane dipeptidase activity; IDA.
 DR GO; GO:0006760; P:folic acid and derivative metabolism; IDA.
 DR InterPro; IPR003137; PA.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF04389; Peptidase_M28; 1.
 DR Pfam; PF04253; TFR_dimer; 1.
 KM Hydroxylase; Carboxypeptidase; Metalloprotease; Zinc; Aminopeptidase;
 KM Dipeptidase; Serine protease; Transmembrane; Glycoprotein;
 KM Signal-anchor; Multifunctional enzyme.
 FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 23 44 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
 FT DOMAIN 45 752 EXTRACELLULAR (POTENTIAL).
 FT ACT_SITE 426 426 NUCLEOPHILE (NINADASE) (BY SIMILARITY).
 FT ACT_SITE 630 630 CHARGE RELAY SYSTEM (POTENTIAL).
 FT ACT_SITE 668 668 CHARGE RELAY SYSTEM (POTENTIAL).
 FT ACT_SITE 691 691 CHARGE RELAY SYSTEM (POTENTIAL).
 FT METAL 379 379 ZINC 2 (BY SIMILARITY).
 FT METAL 389 389 ZINC 1 (BY SIMILARITY).
 FT METAL 427 427 ZINC 2 (BY SIMILARITY).
 FT METAL 455 455 ZINC 1 (BY SIMILARITY).
 FT METAL 555 555 ZINC 2 (BY SIMILARITY).
 FT CONFLICT 141 141 F -> S (IN REF. 2).
 FT CONFLICT 178 178 F -> Y (IN REF. 2).
 FT CONFLICT 219 219 V -> A (IN REF. 2).
 FT CONFLICT 240 240 G -> A (IN REF. 2).
 FT CONFLICT 287 287 E -> N (IN REF. 2).
 FT CONFLICT 583 583 R -> G (IN REF. 2).
 FT CONFLICT 625 625 E -> K (IN REF. 2).
 FT CONFLICT 728 728 S -> N (IN REF. 2).
 FT CONFLICT 749 749 M -> R (IN REF. 2).
 SQ SEQUENCE 752 AA; 84635 MW; DE8F10946A50926A CRC64;

Query Match 7.2%; Score 174; DB 1; Length 752;
 Best Local Similarity 19.9%; Pred. No. 7.2e-05;
 Matches 105; Conservative 83; Mismatches 170; Indels 170; Gaps 24;

QY 3 FLTAFGVGHVLSGSKAICKNGISKRTPEEKIEIASCGDVAKAIINLAVTGAQNR 62
 DB 34 FLIGFLGFWPIKPNENATGVNHSCKMKEFLHEIKAE-----NIKKFLYNF----- 79
 QY 63 SYERLALVDTVGRLSGSK-NLEKAIQIMYQNIQOGGLEKVNHEPRIPMERGESSAV 121
 DB 80 -----TRTHLGLGTQNNFLAKOI-----HDKMKEFGJDLVELSLHYD-----V 117
 QY 122 ML-----EPRIHKIAILGL-GSSI-----GTPREGITAEVVLVVTSPDELGRASEARGKI 170
 DB 118 LLSVNTHTHPYISITINEDNEIFKTSLSBPFGYENISDVVPYSAFSPQGT-PSGLD 176
 QY 171 VVTVQPIYINISRT-----VOYRTOGAVEAAKVGAL-ASLIRSV-----ASFSTYS 214
 DB 177 V-----FVNVARTEDFPKLEREMKISCGSKIVIRYGVKVRGNVKNVQLAGAKGMILYS 231

QY 215 PHT-----GIOEYOD-----GV 226
 DB 232 DPADYFVPGVKSIPDGMNLPBGGVGNVNLNAGDPLTPGVPAHEHAYRHELTAVGL 291
 QY 227 PKIPIACTIVEDAEKMSRNASH-----GIKI-----VIOLKGA 260
 DB 292 PSIPVPIGVDAAQ-----KLEHMGCPAPDPSSWKGGLKPYVNGPGFAGNFSTQKYMMH 348
 QY 261 KTYPD-TDSFNTVAELTSGKYRPOVVLVNSGLDMSVDVQGMDDGAGFISWE---ALSL 316
 DB 349 HSTYKTVRIYVNLVIGTLKGALEPDRVYILGSHRAMWF--GDIPOGSAVVAHVHSFGT 406
 QY 317 IKDLGLRPRKTLRLVLTAEEOGCVAFQYQHLKKNISVNSVLSMESDAGTFLPTGLQFT 376
 DB 407 LKKKGRRPRRTILFASMDAEFFLLGSTWAEH-----SLLDGKGVAYINADSSIE 459
 QY 377 GSEKAPAIKEVMSLLQPLNTIQTVLSHGECTDINFIQACVPASLID 424
 DB 460 GNYTLAVDCTPLMYSL-VVNLTKELQSPD-----EGFEKSKSYD 497

RESULT 5
 GCP2_ARATH STANDARD; PRT; 705 AA.
 AC QPM158;
 DT 28-FEB-2003 (Ref. 41, Created)
 DT 28-FEB-2003 (Ref. 41, Last sequence update)
 DT 28-FEB-2003 (Ref. 41, Last annotation update)
 DE Probable arginase carboxypeptidase II (EC 3.4.17.21).
 GN AMPI OR AT13654720 OR T5N23 80.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta; TISSUE=slilques;
 RX MEDLINE=21434423; PubMed=11549767;
 RA Hellmwell C.A., Chin-Ackins A.N., Wilson I.W., Chapple R.,
 RT Dennis E.S., Chaudhury A.;
 RT "The Arabidopsis AMPI gene encodes a putative glutamate
 carboxypeptidase";
 RL Plant Cell 13:2115-2125(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016720; PubMed=11130713;
 RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseid M.,
 RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermayer B.,
 RA Delaeny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
 RA De Simone V., Choisme N., Artiguenave F., Robert C., Bretter P.,
 RA Wincker P., Catolico L., Weissenbach J., Saurin W., Queller F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wumlbach E., Drzonsek H., Erfle H., Jordan R., Bangert S.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simonati B.,
 RA Conrad A., Hornischer K., Kauer G., Loehart T.H., Nordstiek G.,
 RA Reichelt J., Scharte M., Schoen O., Barynes M., Terol J., Clement J.,
 RA Navarro P., Collado C., Perez-Perez A., Otenmaeider B., Duchemin D.,
 RA Cooke R., Landie M., Berger-Llauró C., Punelle B., Masny D.,
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Collet A., Casacuberta E.,
 RA Montfort A., Argitrou A., Flores M., Lisuori R., Vitale D.,
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walter A., Utecherback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Maiti R., Mu D., Peterson J., Van Aken S.,
 RA Pai G., Miltsober J., Sellers P., Gill J.E., Feldblum T.V.,
 RA Preuss D., Lin X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Frazer C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasamoto S., Kimura T., Ideawa K., Kawashima K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Muraki A.,
 RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,

QY	22	EYDGV-----	-----	PKIPACTVDAEMSMASHG-----	249
Db	251	PGMVGVGSEKSLDELVTTRFFRPIPSLISLRNAEIT--	-----	LASLGARAPLEKRNISGRV	308
QY	250	-----	-----	IKIVQLKMGAKTYPPDTSFNTVAETGSKYEQVVLVSGHLSWDVG	297
Db	309	GPQGRPGPRMVIINMTFOGEMMKIN-----	-----	NVVVITRGEADRVVILIGNRDAMT--	361
QY	298	QGAMDDGGAF----	-----	ISWEASLIDGLRPRTLRVLVLTAEQGVGARFOYQLHKVN	353
Db	362	YGAVDENSSTSLALLISRRFALLTKS--GMRPRRTILCSMPAEFPGMIGSTWIEBNVLN	-----	-----	420
QY	354	IS-----	-----	NSLWMSDAGTFLPRGLQFTGSEKARAIMBVMSLD-----	393
Db	421	LGASAVAYLVNDCAVQG--SGFPAGATFPOLDG-----	-----	LLVDVLKVDPPDPAVGLTBEET	473
QY	394	--PLNTQVLSHGEGDINFVIOAGVNPASLIDDLV--	-----	KYFFPHSHGDTMTVA--	445
Db	474	KQNMNIIOQLSRVSDGDFSGFLHMGIPSI----	-----	DMYVGADVPVYHTAFDSYDMMINHADP	529
QY	446	K-QMNVV-AAWVAVVSYYVADMEEMLP	470	-----	-----
Db	530	LFRHVAMAGIMGLIGILLAD-EPLIP	555	-----	-----
RESULT 6					
YVAD	BACSU	STANDARD;	PRT;	455	AA.
AC	P25152;				
DT	01-MAY-1992 (Rel. 22, Last sequence update)				
DT	01-FEB-1995 (Rel. 31, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Hypothetical peptidase ywad precursor (BC 3.4.11.-)				
GN	YVAD OR IPA-8R.				
OC	Bacillus subtilis.				
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_TaxID=1423;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=168;				
RX	MEDLINE=98044033; PubMed=9384377;				
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,				
RA	Azevedo V., Bertero M.G., Bessieres P., Bojotin A., Borchert S.,				
RA	Bourlles R., Bourster L., Brans A., Braun M., Brignell S.C., Bron S.,				
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Catter N.M.,				
RA	Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,				
RA	Denizot F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T.,				
RA	Enlin C.D., Errington J., Fabret C., Ferrari E., Foulger D.,				
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,				
RA	Ghm S.Y., Glaeser P., Gofeau A., Goldlighty E.J., Grandi G.,				
RA	Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,				
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,				
RA	Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,				
RA	Kobayashi Y., Koelter P., Koningerfein G., Krogh S., Kumano M.,				
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,				
RA	Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,				
RA	Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,				
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,				
RA	Patro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,				
RA	Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,				
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,				
RA	Sato T., Scanlan E., Schleich S., Schroter R., Scofield F.,				

RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Solido B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terasita P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vaasartti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*." ;
 RT Nature 390:249-256(1997).
 RL [3]
 RN SEQUENCE OF 1-68 FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=92216127; PubMed=1806041;
 RA Glaser P., Kunst F., Debarbouille M., Vertes A., Danchin A.,
 RA Dedonder R.;
 RT "A gene encoding a tyrosine tRNA synthetase is located near sacs in
 RT *Bacillus subtilis*." ;
 RL DNA Seq. 1:251-261(1991).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X73124; CAAS1564.1; -;
 DR EMBL: 299123; CAB15873.1; -;
 DR EMBL: X52480; CA36725.1; -;
 DR PIR: S39663; S39663.
 DR MEROPS: M28. UPB.;
 DR Subtilist; BG10554; YWAD.
 DR InterPro: IPR003137; PA.
 DR Pfam: PF02225; PA; 1.
 DR Pfam; PF04389; Peptidase M28; 1.
 KM Hypothetical protein; Hydrolase; Aminopeptidase; Metalloprotease;
 KW Zinc, Signal, Complete proteome.
 FT SIGNAL 1 31 POTENTIAL.
 FT CHAIN 32 455 HYPOTHETICAL PEPTIDASE YWAD.
 FT SEQUENCE 455 AA; 49450 MW; 895E66EB0CCB18 CRC64;
 SQ
 Query March 7.0%; Score 169; DB 1; Length 455;
 Best Local Similarity 23.2%; Pred. No. 8.2e-05;
 Matches 100; Conservative 83; Mismatches 174; Indels 74; Gaps 22;
 QY 58 KAQRSYERLALVDVTPRLSGSKNLEKAIQIMYQNLQDGLKHLPEVPIRPHMERGE 117
 DB 44 KAKH-AISTISQLEAIGPRIAGTAEEKSALLIASSMRKLD-VKQGRNITDRLEGT 101
 QY 118 ESAVMBPRHKIALILGSSIGTPREGITAEVLVTSFDELQRRASARKIVVNOFY 177
 DB 102 LSSAGRD-----ILQAAAGSAPTEBQGLTAP-LYNAGLGYQXFTADAKKIMLSRGD 155
 QY 178 INYERTQYRTQGAWEAKVAKALASLI-RSAVSFIYSPHIGIDYQGVAKIPACTIV 236
 DB 156 LTY-----YEKAKNAEA--GAKAVILYNNESESLVPMPLNSGNV-----GIPVGIKK 203
 QY 237 EDAMEMRMASHGKIKIVQLKMGAKTYPDTSFNTVAITGSKYPE-----QVVLVSGHL 291
 DB 204 EDGALVQOKKATYKL-----KAFNQNSQN-----ITIKKPKNIKHPDIYVTAHY 251
 QY 292 DSWDVGGAMDGGGAPISWEALSLIKDLGAPKRTTALVMTAEQGVAFQY---- 347
 DB 252 DSVFPGGANDGSGTSMLEMARVTK--SVPSDEIRFIKFAEELGLSSHYVDHS 309
 QY 348 --QLHKVNSYSLVM-----ESDAGTLPF--GLQFGSKAKAIMEV---MSLLQF 394
 DB 310 EKEIKRESEV-NFNLDWGTSEWKSSELYVNTLDGNSNYWSSSRPAAEKIFDSLSLTQG 368
 QY 395 LNTQVLSHGEGTD-INF-WIQAGVPGASLLDLYKYFFPHSHSDTWTVMDPKQMN--- 449

DB 369 GSSDHPFPHAGIDISANFIW---GDPETEVEPWW-----HTPEDSIEHISXRLOAG 419
 QY 450 --VAAAWNAV 458
 DB 420 DLVTAAYEAV 430
 RESULT 7
 FOHL_PIG STANDARD; PRT; 751 AA.
 AC 077564;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glutamate carboxypeptidase II (EC 3.4.17.21) (Membrane glutamate
 DE carboxypeptidase) (MGC) (N-acetylated-alpha-linked acidic dipeptidase
 DE I) (NALADase I) (Pteroylpolyl-gamma-glutamate carboxypeptidase)
 DE (Folylpoly-gamma-glutamate carboxypeptidase) (FGCP) (Folate hydrolase
 DE I) (Prostate-specific membrane antigen homolog).
 GN FOHL OR NALADI.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 200-210 AND 471-483.
 RC TISSUE=jejunal mucosa;
 RX MEDLINE=98352082; PubMed=9685395;
 RA Halsted C.H., Ling E.-H., Luthi-Carter R., Villanueva J.A.,
 RA Gardner J.M., Coyle J.T.;
 RT "Polylpoly-gamma-glutamate carboxypeptidase from pig jejunum.
 RT Molecular characterization and relation to glutamate carboxypeptidase
 RT II." ;
 RL J. Biol. Chem. 273:20417-20424 (1998).
 RN [2]
 RP ERRATUM.
 RA Halsted C.H., Ling E.-H., Luthi-Carter R., Villanueva J.A.,
 RA Gardner J.M., Coyle J.T.;
 RL J. Biol. Chem. 275:30746-30746 (2000).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=86085936; PubMed=2867095;
 RA Chandler C.J., Wang T.-T., Halsted C.H.;
 RT "Pteroylpolylglutamate hydrolase from human jejunal brush borders.
 RT Purification and characterization." ;
 RL J. Biol. Chem. 261:928-933 (1986).
 CC -1- FUNCTION: Has both folate hydrolase and N-acetylated-alpha-linked-
 CC acidic dipeptidase (NALADase) activity. Has a preference for tri-
 CC alpha-glutamate peptides (By similarity). In the intestine,
 CC required for the uptake of folate. In the brain, modulates
 CC excitatory neurotransmission through the hydrolysis of the
 CC neuropeptide, N-acetylaspartylglutamate (NAAG), thereby releasing
 CC glutamate. Maximal activity at pH 6.0.
 CC -1- FUNCTION: Also exhibits a dipeptidyl-peptidase IV type activity
 CC (By similarity). In vitro, cleaves Gly-Pro-AMC (By similarity).
 CC CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal
 CC glutamyl residue, typically from Ac-Asp-Glu or Pteroyl-gamma-
 CC -1- COFACTOR: Zinc; Binds two ions per subunit. Required for NALADase
 CC activity.
 CC -1- ENZYME REGULATION: The NALADase activity is inhibited by
 CC quiescinal acid, beta-NAG and 2-(phosphonomethyl) pentanedioic
 CC acid (PMPA). Ethanol ingestion decreases the folate hydrolase
 CC activity by 50%.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Plasma membrane
 CC (By similarity).
 CC -1- TISSUE SPECIFICITY: High expression in the duodenum and in the
 CC jejunum brush-border membrane. Weak expression in kidney.
 CC -1- DOMAIN: The NALADase activity is found in the central region, the
 CC dipeptidyl peptidase IV type activity in the C-terminal.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

RT acidic dipeptidase (NAALDase).";
 RL J. Pharmacol. Exp. Ther. 286:1020-1025 (1998).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM PSMA-1), AND CHARACTERIZATION.
 RC TISSUE=Prostate;
 RX MEDLINE=99185063; PubMed=10085079;
 RA Pangalos M.N., Neels J.-M., Somers M., Verhasselt P., Bakkens M.,
 van der Helm L., Fraiponts E., Ashton D., Gordon R.D.;
 RT "Isolation and expression of novel human glutamate carboxypeptidases
 with N-acetylated alpha-linked acidic dipeptidase and dipeptidyl
 peptidase IV activity.";
 RL J. Biol. Chem. 274:8470-8483 (1999).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORMS PSMA-1 AND PSMA-2), AND VARIANT HHC
 RP TYR-475.
 RC TISSUE=Jejunum;
 RX MEDLINE=20545101; PubMed=11092759;
 RA Devlin A.M., Ling E.-H., Pearson J.M., Fernando S., Clarke R.,
 Smith A.D., Halsted C.H.;
 RT "Glutamate carboxypeptidase II: a polymorphism associated with lower
 levels of serum folate and hyperhomocysteinemia.";
 RL Hum. Mol. Genet. 9:2837-2844 (2000).
 RN [8]
 RP SEQUENCE FROM N.A. (ISOFORM PSMA-5).
 RA Peace D.J., Zhang Y., Holt G., Ferrer K.T., Heller M., Sosman J.A.,
 Xue B.H.;
 RT "Identification of three novel splice variants of prostate-specific
 membrane antigen.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS PSMA-3 AND PSMA-4).
 RA Lipold S.E., Criley S.C., Coffey D.S.;
 RT "Alternative Splicing of the prostate-specific membrane antigen.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE OF 60-74, AND SUBCELLULAR LOCATION.
 RC TISSUE=Prostatic carcinoma;
 RX MEDLINE=99025849; PubMed=9809977;
 RA Grauer I.S., Lawler K.D., Marignac J.L., Kumar A., Geol A.S.,
 Wolfert R.L.;
 RT "Identification, purification, and subcellular localization of
 prostate-specific membrane antigen PSM protein in the LNCaP prostatic
 carcinoma cell line.";
 RL Cancer Res. 58:4787-4789 (1998).
 RN [11]
 RP ALTERNATIVE SPLICING.
 RA Bodega T., She D., Turi T., Wroblewska B., Neale J.R.;
 RT "Molecular cloning of alternatively spliced variants of the peptidase
 against N-acetylaspartylglutamate (NAAG) from human and rat nervous
 systems.";
 RL Abstr. Soc. Neurosci. 24:579-579 (1998).
 RN [12]
 RP CHARACTERIZATION.
 RX MEDLINE=98288196; PubMed=9622670;
 RA Luthi-Carter R., Barczak A.K., Speno H.D., Coyle J.T.;
 RT "Hydrolysis of the neuropeptide N-acetylaspartylglutamate (NAAG) by
 cloned human glutamate carboxypeptidase II.";
 RL Brain Res. 795:341-348 (1998).
 RN [13]
 RP DOMAIN STRUCTURE.
 RX MEDLINE=97330810; PubMed=9187245;
 RA Rawlings N.D., Barrett A.J.;
 RT "Structure of membrane glutamate carboxypeptidase.";
 RL Biochim. Biophys. Acta 1339:247-252 (1997).
 RN [14]
 RP MUTAGENESIS.
 RX MEDLINE=99102317; PubMed=9882712;
 RA Speno H.S., Luthi-Carter R., Macias W.L., Valentine S.L.,
 Joshi A.R.T., Coyle J.T.;
 RT "Site-directed mutagenesis of predicted active site residues in
 glutamate carboxypeptidase II.";
 RL Mol. Pharmacol. 55:179-185 (1999).
 CC -1- FUNCTION: Has both folate hydroxylase and N-acetylated-alpha-linked-

CC acidic dipeptidase (NAALDase) activity. Has a preference for tri-
 CC alpha-glutamate peptides. In the intestine, required for the
 CC uptake of folate. In the brain, modulates excitatory
 CC neurotransmission through the hydrolysis of the neuropeptide, N-
 CC acetylaspartylglutamate (NAAG), thereby releasing glutamate. Stable
 CC at pH greater than 6.5. Isoforms PSM-4 and PSM-5 would appear to
 CC be physiologically irrelevant. Involved in prostate tumor
 CC progression.
 CC -1- FUNCTION: Also exhibits a dipeptidyl-peptidase IV type activity.
 CC In vitro, cleaves Gly-Pro-AMC.
 CC -1- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal
 CC glutamyl residue, typically from Ac-Asp-Glu or pteroyl-gamma.
 CC COFACTOR: Zinc; Binds two ions per subunit. Required for NAALDase
 CC activity.
 CC -1- ENZYME REGULATION: The NAALDase activity is inhibited by beta-
 CC NAAG, quaggaic acid, 2-(phosphonomethyl) pentanedioic acid
 CC (FMPA) and EDTA. Activated by cobalt.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Plasma membrane.
 CC The PSMA isoform is cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=6;
 CC Comment=Experimental confirmation may be lacking for some
 CC isoforms;
 CC Name=PSMA-1;
 CC IsoId=Q04609-1; Sequence=Displayed;
 CC Name=PSMA-2;
 CC IsoId=Q04609-2; Sequence=VSP_005341;
 CC Name=PSMA-3;
 CC IsoId=Q04609-3; Sequence=VSP_005342;
 CC Name=PSMA-4;
 CC IsoId=Q04609-4; Sequence=VSP_005339, VSP_005340;
 CC Name=PSMA-5;
 CC IsoId=Q04609-5; Sequence=VSP_005337, VSP_005338;
 CC Name=PSMA;
 CC IsoId=Q04609-6; Sequence=VSP_005336;
 CC -1- TISSUE SPECIFICITY: Highly expressed in prostate epithelium. Also
 CC expressed, in the small intestine, brain, kidney, liver, spleen,
 CC colon and the capillary endothelium of a variety of tumors.
 CC Expressed specifically in jejunum brush border membranes. In the
 CC brain, highly expressed in the ventral striatum and brain stem. In
 CC the prostate, the PSMA' cytosolic isoform is the most abundant
 CC form in normal tissue, the membrane-bound PSMA-1 form in primary
 CC prostate tumors. The PSMA-2 isoform also found in normal prostate
 CC as well as in brain and liver.
 CC -1- INDUCTION: In the prostate, up-regulated in response to androgen
 CC deprivation.
 CC -1- DOMAIN: The NAALDase activity is found in the central region, the
 CC dipeptidyl peptidase IV type activity in the C-terminal.
 CC -1- PTM: The first two amino acids at the N-terminus of isoform PSMA'
 CC appear to be cleaved by limited proteolysis.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: Defects in FOLH1 may be a cause of hyperhomocysteinemia
 CC (HHCI), a condition associated with increased risk of
 CC cardiovascular disease, neural tube defects, and cognitive
 CC deficits.
 CC -1- MISCELLANEOUS: PSMA is used as a diagnostic and prognostic
 CC indicator of prostate cancer, and as a possible marker for various
 CC neurological disorders such as schizophrenia, Alzheimer's disease
 CC and Huntington's disease.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; W99487; AAA60209.1; -;
 CC EMBL; S76978; AAB33750.2; -;
 CC EMBL; AF007544; AAC83972.1; -;


```

Db 450 NISLHVDTVPLEHVAIVEASQVENSRSKRERKRLTYDWMKVPDCKAGVFKIRVPG 509
Qy 405 EGTD-----INFWIAGVPGASLDDLYKFFPHSHGDT-----MTWMDPK 446
Db 510 GGSDDHAFPLNF--AGVP-----VINFTKNTWDTPLFHTMYETPPSNHLLDPTD 559
Qy 447 QMNVAAAV---WAVSVYVADMEEMLP 470
Db 560 NLSVHKAIGGYWMAELAKTFND-DVILP 585

RESULT 10.
TFRL FELCA STANDARD; PRT; 769 AA.
AC Q9MYZ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Transferrin receptor protein 1 (TfR1) (TR) (TfR) (Tftr).
GN TRC.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21165299; PubMed=11264378;
RA Parker J.S.L., Murphy W.J., Wang D., O'Brien S.J., Parrish C.R.;
RT "Canine and feline parvoviruses can use human or feline transferrin
RL J. Virol. 75:3896-3902(2001).
CC -1- FUNCTION: Cellular uptake of iron occurs via receptor-mediated
CC endocytosis of ligand-occupied transferrin receptor into
CC specialized endosomes. Endosomal acidification leads to iron
CC release. The apotransferrin-receptor complex is then recycled to
CC the cell surface with a return to neutral pH and the concomitant
CC loss of affinity of apotransferrin for its receptor. Transferrin
CC receptor is necessary for development of erythrocytes and the
CC nervous system (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Binds one transferrin
CC molecule per polypeptide chain (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -1- PTM: N- and O-glycosylated, phosphorylated and palmitoylated (By
CC similarity).
CC -1- MISCELLANEOUS: Canine and feline parvoviruses bind human and
CC feline transferrin receptors and use these receptors to enter and
CC infect cells.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR HSMBL; AF276984; AAF81908.1; -.
DR HSBP; P02786; 1CX8.
DR InterPro; IPR003137; PA.
DR Pfam; PF02225; PA; 1.
DR Pfam; PF04389; Peptidase_M28; 1.
DR Pfam; PF04253; TFR_dimer; 1.
KW Transmembrane; Glycoprotein; Receptor; Lipoprotein; Palmitate;
KW Signal-anchor; Endocytosis; Phosphorylation.
KW DOMAIN 1 70 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 71 91 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 92 769 (POTENTIAL).
FT DOMAIN 578 769 EXTRACELLULAR (POTENTIAL).
FT SITE 20 23 LIGAND-BINDING (BY SIMILARITY).
FT ENDOCYTOSIS SIGNAL.

```

```

FT SITE 61 64 STOP-TRANSFER SEQUENCE.
FT SITE 655 657 CELL ATTACHMENT SITE (POTENTIAL).
FT LIPID 70 70 PALMITATE (BY SIMILARITY).
FT MOD RES 24 24 PHOSPHORYLATION (BY SIMILARITY).
FT DISULFID 92 92 INTERCHAIN (BY SIMILARITY).
FT DISULFID 101 101 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 107 107 O-LINKED (GLNAC. . .) (BY SIMILARITY).
FT CARBOHYD 260 260 N-LINKED (GLCNAAC. . .) (BY SIMILARITY).
FT CARBOHYD 326 326 N-LINKED (GLCNAAC. . .) (BY SIMILARITY).
FT CARBOHYD 731 731 N-LINKED (GLCNAAC. . .) (POTENTIAL).
FT CARBOHYD 736 736 N-LINKED (GLCNAAC. . .) (BY SIMILARITY).
SQ SEQUENCE 769 AA; 86270 MW; 47EA4FCA18A46D7 CRC64;

Query Match 6.4%; Score 153.5; DB 1; Length 769;
Best Local Similarity 21.0%; Pred. No. 0.0023;
Matches 101; Conservative 65; Mismatches 177; Indels 139; Gaps 24;

Qy 76 PRLSGSKNLEKAIQIMYQNLQDGLKRV---HLEPRIPHWERGESAVMLBPRIRKIA 131
Db 163 PREAGSQKDSLAFFIENRPRELQSKAMWDEHVKVQV---KGSAS-----NSVT 210
Qy 132 ILGLGSS---IGTPPEGITA--EVLVYT-----SPDELQRASEARGKIVYN 174
Db 211 IVGTNSGMVYLVSPEGYVAYSXAATVGRVHANFGTKDFENLN---SPVNGSLIVR 267
Qy 175 QPVYINSGRTQYVRQGAWEAKVGALASLI-RSVAFSI-----YSP 215
Db 268 AGKITPFAKV-----ANAEFRALGVLYIMDAKAPITNAELPFGGHAHLTGDPYTP 320
Qy 216 -----HTGIQEOYQ-DGVPIKTPACTIVEDAEMMSRWASHGIKIVILKNGAKTYPDTDS 268
Db 321 GPFSPNTQPPPSQSSSLPNIPOGISRANAE-----KLFGMWEDGCSAMETDS 370
Qy 269 -----PVTVAETGSKPEQVYLVSHGLSDVDGQGMMD 303
Db 371 SCRLETSRMNNVYKLSVNNVLKEIRIFVFGVIGFEEPDHYVVGQORDAM--GPGAKS 428
Qy 304 GGAFFISWEALSLIKDL-----GLRPKRTLLVMTAEQGVGAFQYQY--LHKVNI SNY 357
Db 429 SVGTALLLELARIISDMVLKGGFKPSSIVPASKSADPGAVGTEMLBGLYSLHLKAF 488
Qy 358 SLVWESDAGTFLPTGLQFTGSEKARALMEBVM-----SLAQPLN-ITVYLSHG-E 405
Db 489 TYI-NLDKAVLGRNSFVKSASPLLSLIEKVMQVKKHPVGGSLYRSNNINKYKESLD 547
Qy 406 GTDINFWIQGVPGASLI---DDLKYKFFPHSHGDTMTWMDP-----KQMNVAAYAV 457
Db 548 NAAFPFLVAGIYAVSFCFCEDTDYPYL-----GITMDVYEKLQKVPQILNKMAARAAB 601
Qy 458 VS 459
Db 602 VA 603

RESULT 11
AMPX_VIBPR STANDARD; PRT; 504 AA.
AC Q01693;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Bacterial leucyl aminopeptidase precursor (EC 3.4.11.10).
OS Vibrio proteolyticus (seromona proteolytica).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=671;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 107-136 AND 233-405.
RC STRAIN=ATCC 15338;
RX MEDLINE=92329552; PubMed=1627651;
RA van Heeke G., Denslow S., Watkins J., Wilson K., Wagner F.;
RT "Cloning and nucleotide sequence of the vibrio proteolyticus
aminopeptidase gene.";

```

Query Match	Similarity	Score	DB 1	Length	504
Beet local	23.0%	Pred. No. 0.0034			
Matches	70	Conservative	47	Mismatches	110
				Indels	77
				Gaps	12
192	VEAAKVGALALISVSASF--IYSPHTGIGQ-----EYODGVKPIPTACITVEDAEEM	242			
123	VDAQGI---TGTISLESFETNRFTYTTTSGAQAQADWIAEWMALSLAPNASV-----	171			
243	SRMASHGIIKIVIQLMGAKTYPDTDSFNTVAETGSKTPEQVVLVSGHLS-----W	294			
172	-KQVSHS-----GYNQKSVVMVTITGSEAPDEWIVIGCHLDSTIGSHTNEQ	215			
295	DVGGGAMDDGGAFISWALSLIDKLGAPRRTLRVLVMTAEQGVG---AFQYQQLH	350			
216	SVAPPADADGAGIAAVTEVIVLISNNPQPKRSIAFMAYAAEEVGLRSGQDLANQYKSEG	275			
351	KVNISNYSVLWESDAGT-----FLPTGLQFTGSEKARAIMEEVMSILQPLNITQVLSHGE	405			
276	KNVVSLALQDMTNTYKGSADQVVFITDITDSNFTQVLTQLMDEYL-----PSLITVGF	326			
406	GT-----DINFWIOAGVPGASLLDDLK-YFFHSHGDTWTWDP-----KQNVV	450			
327	DTCGVACGDHSHMHNAGYPAAMPFSKFNQVNPRIHTTQDPLANSDDPTGSHAKKFTQLGL	386			
451	AAAV 454				

Db	387	AYAI	390
RESULT 12			
ID	TFRL_RAT	STANDARD;	PRT; 622 AA.
AC	O99376;		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	transferrin receptor protein 1 (TFRL) (TR) (Tfr) (Fragment).		
GN	TFRC OR TFRF.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI TaxID=10116;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Testis;		
RX	MEDLINE=91125359; PubMed=2126342;		
RA	Roberts K.P., Griswold M.D.,		
RT	"Characterization of rat transferrin receptor cDNA: the regulation of		
RT	transferrin receptor mRNA in testes and in Sertoli cells in culture.";		
RL	Mol. Endocrinol. 4:531-542(1990).		
CC	-1- PUNCTATION: Cellular uptake of iron occurs via receptor-mediated		
CC	endocytosis of ligand-occupied transferrin receptor into		
CC	specialized endosomes. Endosomal acidification leads to iron		
CC	release. The apotransferrin-receptor complex is then recycled to		
CC	the cell surface with a return to neutral pH and the concomitant		
CC	loss of affinity of apotransferrin for its receptor. Transferrin		
CC	receptor is necessary for development of erythrocytes and the		
CC	nervous system (By similarity).		
CC	-1- SUBUNIT: Homodimer; disulfide-linked. Binds one transferrin		
CC	molecule per polypeptide chain (By similarity).		
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).		
CC	-1- TISSUE SPECIFICITY: In testis, expressed in Sertoli cells,		
CC	peritubular myoid cells and in germinal cells. Highest levels in		
CC	Sertoli cells.		
CC	-1- PTM: N- and O-glycosylated, phosphorylated and palmitoylated (By		
CC	similarity).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch).		
CC	or send an email to license@isb-sib.ch .		
CC	-----		
DR	EMBL; M58040; AAA42273.1; -.		
DR	PIR; A34549; A34549.		
DR	HSSP; P02786; 1CX8.		
DR	InterPro; IPR003137; PA.		
DR	Pfam; PF02225; PA; 1.		
DR	Pfam; PF04389; Peptidase_M28; 1.		
DR	Pfam; PF04253; TFR_dimer; 1.		
KW	Transmembrane; Glycoprotein; Receptor; Lipoprotein; Palmitate;		
KW	Signal-anchor; Endocytosis; Phosphorylation.		
FT	NON_TER	1	
FT	DOMAIN	431	622
FT	SITE	508	510
FT	CARBOHYD	109	109
FT	CARBOHYD	113	113
FT	CARBOHYD	179	179
FT	CARBOHYD	584	584
FT	CARBOHYD	589	589
SQ	SEQUENCE	622 AA;	70152 MW; 831EFCLDE55703 CRC64;
Query March	5.9%;	Score 143.5;	DB 1; Length 622;
Best local	Similarity	19.4%;	Pred. No. 0.0088;
Matches	90;	Conservative	61; Mismatches 158; Indels 155; Gaps 21.

[illegible]

[illegible]

FT CONFLICT 248 248 S -> L (IN REF. 2).
 FT CONFLICT 287 287 A -> V (IN REF. 2).
 FT CONFLICT 595 595 K -> E (IN REF. 1).
 SO SEQUENCE 798 AA; 88402 MW; FA6161FE3FFFA4 CRC64;

Query Match 5.9%; Score 142; DB 1; Length 798;
 Best Local Similarity 18.4%; Pred. No. 0.016;
 Matches 104; Conservative 69; Mismatches 190; Indels 202; Gaps 22;

QY 9 FGGHLLSCSGKAIKCKGKISKTEKEERIASGCD--VAKAIIN----- 52
 DB 91 FTGAFLGYVAFVFGSCQ-----ACGDSVLVDEEDVNEDESGRTLLYMSD 134
 QY 53 -----LAVYGRKONSRYERIALLVTVGPRSLGSGNLEKAIIVYONLQODGLEKY----- 103
 DB 135 LQAMLRRLGEGRMEDTIRLTSLR-----RVAGSKAMATLVODIDKLRSOKLDHVMTDT 190
 QY 104 HLEPVRI-----HWERG-----ESSAVMLEPRIH-----KIALIGSSIG 140
 DB 191 HVGGLQFPDPAHANTLHWVDADGVSQEQLEDPEDVYCPYSATGNATGKLVYAHYGRS-- 248
 QY 141 TPPEGTAEVLVYTSFDELQRRASARCKIYVYNOPIYNSKTYQRTQGAVERAKVGL 200
 DB 249 -----EDLODLKAKGVELAGSLHLVRVGLTSFAQKV-----AVAQDFGAQ 288
 QY 201 ASLIRSVASFISYSPH-TGIOEYO-----DGVPK 228
 DB 289 GVLVYDSDSDSODHPKPKGLSHQVNHVHGLTGDPTTPGPPSPNQTQPPVSSGLPS 348
 QY 229 IPTACTIVEDAEMMSRMAS-----HGKIYIOLKMGAKTYPD-----TD 267
 DB 349 IPAGPISADIDQLRLKLTGPVAPQEMKGLSGSPYRLGPG--PDRLVYNNHVRVSTP 404
 QY 268 SPNTVAETGSKYEPQVVLVSGHLSWDVCGGAMDDGGAFISWALSILIDL--GLRP 324
 DB 405 ISNIFACIEGPAEPHYVIVIGARDAM--GPGAAKSAVGTILLETVRTFSSMVSNGFRP 462
 QY 325 KRTLRVMTAEBOGVGAFOYQLHKVNISYLVMSDSADGTFPLPTGLQFTGSEKANA- 383
 DB 463 RSLFLISWDGDDFGSSVATKEM-----LEGIYSLVHLKAVYVSLNSVLGDKFKPAK 515
 QY 384 -----IMEEVM-----LLOPLNITQVLSHGEGTDIN 410
 DB 516 TSPILVSLIENILKQVDSPNHSGQTLVRYQVALTHPSWDAVYIQLPM-----DSSAYS 568
 QY 411 FWIQGVPGA--SLDDLYKTYFFH 433
 DB 569 FTAFAGVPAVEFSFMEDDRVPFLH 593

RESULT 15
 APX_STRGR STANDARD; PRT; 284 AA.
 ID APX_STRGR
 AC P80561;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Aminopeptidase (EC 3.4.11.-) (SGAP).
 OS Streptomyces griseus.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1911;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96270734; PubMed=8665903;
 RA Maras B., Greenblatt H.M., Shoham G., Spungin-Bialik A., Blumberg S.,
 RA Barra D.,
 RT "Aminopeptidase from Streptomyces griseus: primary structure and
 RT comparison with other zinc-containing aminopeptidases.";
 RL Eur. J. Biochem. 236:843-846 (1996).
 RN [2]
 RP SEQUENCE OF 1-6.
 RX MEDLINE=89338422; PubMed=2503378;

RA Spungin A., Blumberg S.;
 RT "Streptomyces griseus aminopeptidase is a calcium-activated zinc
 RT metalloprotein. Purification and properties of the enzyme.";
 RL Eur. J. Biochem. 185:471-477 (1989).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RX MEDLINE=97201142; PubMed=9048953;
 RA Greenblatt H.M., Almog O., Maras B., Spungin-Bialik A., Barra D.,
 RA Blumberg S., Shoham G.;
 RT "Streptomyces griseus aminopeptidase: X-ray crystallographic
 RT structure at 1.75-A resolution.";
 RL J. Mol. Biol. 265:620-636 (1997).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=93185612; PubMed=8444149;
 RA Ben-Meir D., Spungin A., Ashkenazi R., Blumberg S.;
 RT "Specificity of Streptomyces griseus aminopeptidase and modulation of
 RT activity by divalent metal ion binding and substitution.";
 RL Eur. J. Biochem. 212:107-112 (1993).
 CC -1- FUNCTION: SPECIFIC FOR LARGER HYDROPHOBIC ACIDS, ESPECIALLY
 CC LEUCINE. NO CLEAVAGE OCCURS IF THE NEXT RESIDUE IS PROLINE.
 CC -1- COFACTOR: BINDS 2 ZINC IONS.
 CC -1- ENZYME REGULATION: BINDS A CALCIUM ION WHICH MODULATES THE
 CC ACTIVITY OF THE ENZYME.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=29728; MW ERR=1.0; METHOD=Electrospray.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
 DR PIR; S66427; S66427.
 DR PDB; 1XJO; 01-APR-97.
 DR PDB; 1CP7; 03-MAY-00.
 DR PDB; 1F20; 22-AUG-01.
 DR PDB; 1F2P; 22-AUG-01.
 DR PDB; 1QO3; 03-MAY-00.
 DR MEROPS; M28.003; -.
 DR Pfam; PF04389; Peptidase_M28; 1.
 KW Hydrolase; Aminopeptidase; Zymogen; Zinc; Calcium; 3D-structure.
 FT METAL 85 97 ZINC 1.
 FT METAL 132 132 ZINC 2.
 FT METAL 160 160 ZINC 1.
 FT METAL 247 247 ZINC 2.
 FT DISULFD 245 250
 FT HELIX 6 22
 FT TURN 23 25
 FT STRAND 28 28
 FT TURN 29 30
 FT HELIX 32 47
 FT TURN 48 49
 FT STRAND 51 59
 FT TURN 60 61
 FT STRAND 62 71
 FT STRAND 75 85
 FT STRAND 88 88
 FT TURN 90 91
 FT TURN 95 98
 FT HELIX 99 115
 FT STRAND 116 116
 FT STRAND 121 128
 FT HELIX 131 133
 FT TURN 134 134
 FT HELIX 136 144
 FT HELIX 147 150
 FT TURN 151 152
 FT STRAND 153 159
 FT STRAND 170 171
 FT HELIX 176 189
 FT TURN 190 190
 FT STRAND 194 194
 FT TURN 204 204
 FT HELIX 205 210
 FT TURN 211 212
 FT STRAND 215 219

FT STRAND 225 225
 FT HELIX 228 234
 FT TURN 238 239
 FT STRAND 241 241
 FT TURN 244 247
 FT TURN 249 250
 FT TURN 253 254
 FT HELIX 258 275
 FT TURN 276 276
 SQ SEQUENCE 284 AA; 29722 MW; 671B80F8CASC4CC CRC64;

Query Match 5.7%; Score 138.5; DB 1; Length 284;
 Best Local Similarity 21.6%; Pred. No. 0.0067;
 Matches 63; Conservative 42; Mismatches 113; Indels 73; Gaps 11;

QY 227 PKIPTACITVEDDAMMSRMASHGIKIV-----IQLKMGAKTYPDT----- 266
 DB 2 PDIPLANVKAHLTQLSTIAANNNGNRAHGRPGYKASVDYVAKLDAAGYTTTLOQFTSGG 61
 QY 267 -DSFNTVAEITGSKYPCQVVLVSGHLDSDWDVGOGAMDGGGAFISWEALSLIKDIGLRPK 325
 DB 62 ATGYNLIANNWGGD-PNKVLMAGAHLDVSSGAGINDNGSSAAVLETALAVSRAGYQPD 120
 QY 326 RTLRVLVMTAEEOGGVGA-FQYQOLHKVNIS-----NYSIWESDAGTFLPTGLQFTGS 378
 DB 121 KHLRFAMWGABELGLIGSKFYVNNLPSADRSKLAGYLNFDWIGSPNPGYFV-----Y 172
 QY 379 EKAAIMEEWSLQPLNI-TQVLSHGEG-TDINFMIQAGVPGASL----- 422
 DB 173 DDDPVIKTEPKNYPAGLNVPTEIETEGDGRSDHAPFKNVGVVVGGLFTGAGYTKSAQAQ 232
 QY 423 -----LDDLYKYFFFHSHGDTM-----TWMDPKQMWAAAVMAVVS 459
 DB 233 KMGGTAGQAFPRCY-----HSCDSLSENINDTALDRNSDAAAAHAWTLSS 277

Search completed: December 22, 2003, 11:50:00
 Job time : 32.9104 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: December 22, 2003, 11:37:00 ; Search time 71.493 Seconds
(without alignments)
1703.674 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417
Sequence: 1 MKFLIFAFPGVHLLSLCSG.....AWMAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteria:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2417	100.0	472	4	Q9Y646
2	2402	99.4	472	4	Q8NBZ1
3	2285	94.5	541	4	Q9Y5X6
4	2133	88.2	470	11	Q9WVJ3
5	2117	87.6	472	11	Q9JLV0
6	2116	87.5	472	11	Q9JLY1
7	1907	78.9	433	11	Q70216
8	759.5	31.4	493	5	Q76552
9	701	29.0	140	4	Q9UNM8
10	655	27.1	414	16	Q8EBH7
11	640	26.5	467	16	Q9A5A8
12	623	25.8	139	11	Q9RI14
13	499.5	20.7	472	16	Q8PHB8
14	420	17.4	375	16	Q8P625
15	193.5	8.0	449	2	Q9JEB5
16	182	7.5	536	16	Q9H208

17	176.5	7.3	571	16	Q9A4J1	Q9A4J1 caulobacter
18	169	7.0	393	2	Q82996	Q82996 aeromonas p
19	165	6.8	466	16	Q9K671	Q9K671 bacillus ha
20	165	6.8	719	4	Q8TAY3	Q8TAY3 homo sapien
21	163	6.7	1483	5	Q8PGK5	Q8PGK5 xanthomonas
22	163	6.7	501	16	Q9KLD3	Q9KLD3 caenorhabdi
23	158.5	6.6	485	2	Q53737	Q53737 streptomyce
24	157.5	6.5	433	17	Q9HRR8	Q9HRR8 halobacteri
25	156	6.5	501	2	P61527	P61527 vibrio chol
26	155.5	6.4	536	16	Q8PFH7	Q8PFH7 xanthomonas
27	151	6.2	529	16	Q9PF58	Q9PF58 xylella fas
28	150.5	6.2	606	10	Q94JH3	Q94JH3 oryza sativ
29	149.5	6.2	609	2	Q8VU54	Q8VU54 alteromonas
30	149.5	6.2	500	16	P96264	P96264 mycobacteri
31	149	6.2	441	17	Q9HNP7	Q9HNP7 halobacteri
32	147.5	6.1	430	16	Q9A3U5	Q9A3U5 caulobacter
33	146	6.0	536	16	Q8P326	Q8P326 xanthomonas
34	145	6.0	488	17	Q9UXB4	Q9UXB4 sulfolobus
35	140	5.8	374	3	Q04033	Q04033 saccharomyc
36	138.5	5.7	594	16	Q8PFF6	Q8PFF6 xanthomonas
37	138.5	5.7	621	10	Q94TH4	Q94TH4 oryza sativ
38	137.5	5.7	324	16	Q9F2X2	Q9F2X2 streptomyce
39	134	5.5	768	6	Q8H2V3	Q8H2V3 sus scrofa
40	134	5.5	384	3	Q8WZH8	Q8WZH8 agaricus bi
41	133.5	5.5	469	16	Q99ZU1	Q99ZU1 streptococc
42	130	5.4	469	16	Q8K716	Q8K716 streptococc
43	130	5.4	280	10	Q9X428	Q9X428 arabidopsis
44	129	5.3	469	16	Q8P162	Q8P162 streptococc

ALIGNMENTS

RESULT 1

Q9Y646 PRELIMINARY; PRT; 472 AA.
ID Q9Y646
AC Q9Y646;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Amino peptidase (Plasma glutamate carboxypeptidase).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Liu C.H., Lin B.Y., Chang L.Y.;
RT "Cloning of the human aminopeptidase gene."
RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strauberg R.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF107834; AAD3214.1; -;
DR EMBL; BC020689; AAH20689.1; -;
DR MEROPS; M28.014; -;
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA.1.
KW Carboxypeptidase.
SQ SEQUENCE 472 AA; 51887 MW; BB6CRD2149E042BF CRC64;

Query Match 100.0%; Score 2417; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 3.6e-176;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEETKEETASCGDVAKAIINLAVYKAQ 60
DB 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEETKEETASCGDVAKAIINLAVYKAQ 60

```
OY 61 NRSYERLALLVDVTGPRLSGKNLEKAIQIMYQNLQODGLEKVLHPVRIIPMERGESBA 120
DB 61 NRSYERLALLVDVTGPRLSGKNLEKAIQIMYQNLQODGLEKVLHPVRIIPMERGESBA 120
OY 121 VMLERIRIKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180
DB 121 VMLERIRIKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180
OY 181 SRTVOYRTOGAVEAKVAGALASLIRSASFISYSPHTIGIYQYODGVPKIPRACTIVEDAE 240
DB 181 SRTVOYRTOGAVEAKVAGALASLIRSASFISYSPHTIGIYQYODGVPKIPRACTIVEDAE 240
OY 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGKYPEQVVLVSGHLSMDVQGA 300
DB 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGKYPEQVVLVSGHLSMDVQGA 300
OY 301 MDDGGAFISWEALSILKDLGRPKRTLRLVMTAEBOGVGAFQYQOLHKVNISNYSLV 360
DB 301 MDDGGAFISWEALSILKDLGRPKRTLRLVMTAEBOGVGAFQYQOLHKVNISNYSLV 360
OY 361 MESDAGTFLPTGLQFTSGSEKARAIIMEEYMSLLQPLNITQVLSHGEGTDINFWIOAGVGA 420
DB 361 MESDAGTFLPTGLQFTSGSEKARAIIMEEYMSLLQPLNITQVLSHGEGTDINFWIOAGVGA 420
OY 421 SLDDLKYKFFPHSHGDTMTVMDPKOMNVAANAVALVSYVADMEMLPRS 472
DB 421 SLDDLKYKFFPHSHGDTMTVMDPKOMNVAANAVALVSYVADMEMLPRS 472
```

RESULT 2

```
O8NBZ1 PRELIMINARY; PRT; 472 AA.
ID Q8NBZ1. PRELIMINARY; PRT; 472 AA.
AC Q8NBZ1. PRELIMINARY; PRT; 472 AA.
BT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein FLJ90651.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCB1_TaxID=9606;
RN NCB1_TaxID=9606;
RP SEQUENCE FROM N.A.
RC Tissue=Placenta;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuho Y., Ota T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK075132; BAC11423.1; -.
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 51784 MW; 83F25996F5DC52F3 CRC64;
```

```
Query Match 99.4%; Score 2402; DB 4; Length 472;
Best Local Similarity 99.4%; Pred. No. 5e-175;
Matches 469; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 MKFLIFAFGGVHLLSLCSGKAICNGISKRTFEIRKEIASCGVAKAIINLAVYGAQ 60
DB 1 MKFLIFAFGGVHLLSLCSGKAICNGISKRTFEIRKEIASCGVAKAIINLAVYGAQ 60
OY 61 NRSYERLALLVDVTGPRLSGKNLEKAIQIMYQNLQODGLEKVLHPVRIIPMERGESBA 120
DB 61 NRSYERLALLVDVTGPRLSGKNLEKAIQIMYQNLQODGLEKVLHPVRIIPMERGESBA 120
OY 121 VMLERIRIKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180
DB 121 VMLERIRIKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180
OY 181 SRTVOYRTOGAVEAKVAGALASLIRSASFISYSPHTIGIYQYODGVPKIPRACTIVEDAE 240
DB 181 SRTVOYRTOGAVEAKVAGALASLIRSASFISYSPHTIGIYQYODGVPKIPRACTIVEDAE 240
```

```
OY 181 SRTVOYRTOGAVEAKVAGALASLIRSASFISYSPHTIGIYQYODGVPKIPRACTIVEDAE 240
DB 181 SRTVOYRTOGAVEAKVAGALASLIRSASFISYSPHTIGIYQYODGVPKIPRACTIVEDAE 240
OY 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGKYPEQVVLVSGHLSMDVQGA 300
DB 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGKYPEQVVLVSGHLSMDVQGA 300
OY 301 MDDGGAFISWEALSILKDLGRPKRTLRLVMTAEBOGVGAFQYQOLHKVNISNYSLV 360
DB 301 MDDGGAFISWEALSILKDLGRPKRTLRLVMTAEBOGVGAFQYQOLHKVNISNYSLV 360
OY 361 MESDAGTFLPTGLQFTSGSEKARAIIMEEYMSLLQPLNITQVLSHGEGTDINFWIOAGVGA 420
DB 361 MESDAGTFLPTGLQFTSGSEKARAIIMEEYMSLLQPLNITQVLSHGEGTDINFWIOAGVGA 420
OY 421 SLDDLKYKFFPHSHGDTMTVMDPKOMNVAANAVALVSYVADMEMLPRS 472
DB 421 SLDDLKYKFFPHSHGDTMTVMDPKOMNVAANAVALVSYVADMEMLPRS 472
```

RESULT 3

```
O9YSX6 PRELIMINARY; PRT; 541 AA.
ID Q9YSX6. PRELIMINARY; PRT; 541 AA.
AC Q9YSX6. PRELIMINARY; PRT; 541 AA.
BT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Blood plasma glutamate carboxypeptidase precursor (EC 3.4.17.21).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCB1_TaxID=9606;
RN NCB1_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=99223495; PubMed=10206990;
RA Gingras R., Richard C., El-Alfy M., Morales C.R., Potier M.,
RA Peshetsky A.V.;
RT "Purification, cDNA cloning, and expression of a new human blood
RT plasma glutamate carboxypeptidase homologous to N-acetyl-aspartyl-
RT alpha-glutamate carboxypeptidase/prolactin-specific membrane antigen.";
RL J. Biol. Chem. 274:11742-11750 (1999).
DR EMBL; AF119386; AAD31418.1; -.
DR MEROPS; M28.014; -.
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
KW Carboxypeptidase; Hydrolyase; Signal.
FT SIGNAL 1 24
FT CHAIN 45 541 BLOOD PLASMA GLUTAMATE CARBOXYPEPTIDASE.
SQ SEQUENCE 541 AA; 59931 MW; 9B138FES7081F68 CRC64;
```

```
Query Match 94.5%; Score 2285; DB 4; Length 541;
Best Local Similarity 99.6%; Pred. No. 5.4e-166;
Matches 445; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 MKFLIFAFGGVHLLSLCSGKAICNGISKRTFEIRKEIASCGVAKAIINLAVYGAQ 60
DB 1 MKFLIFAFGGVHLLSLCSGKAICNGISKRTFEIRKEIASCGVAKAIINLAVYGAQ 60
OY 61 NRSYERLALLVDVTGPRLSGKNLEKAIQIMYQNLQODGLEKVLHPVRIIPMERGESBA 120
DB 61 NRSYERLALLVDVTGPRLSGKNLEKAIQIMYQNLQODGLEKVLHPVRIIPMERGESBA 120
OY 121 VMLERIRIKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180
DB 121 VMLERIRIKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180
OY 181 SRTVOYRTOGAVEAKVAGALASLIRSASFISYSPHTIGIYQYODGVPKIPRACTIVEDAE 240
DB 181 SRTVOYRTOGAVEAKVAGALASLIRSASFISYSPHTIGIYQYODGVPKIPRACTIVEDAE 240
OY 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGKYPEQVVLVSGHLSMDVQGA 300
DB 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGKYPEQVVLVSGHLSMDVQGA 300
```

```

Db 241 MMSHAGHGIKIYVQLKMGAKTYPDTSFNTVAETISGSKYPEQVVLVSGHSDMDVCGA 300
Qy 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEEGGVGAFQYQLHVNINSNLSV 360
Db 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEEGGVGAFQYQLHVNINSNLSV 360
Qy 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLLOPLNTQVLSHGEGDINFWIOAGVGA 420
Db 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLLOPLNTQVLSHGEGDINFWIOAGVGA 420
Qy 421 SLDDLKYYFFHHSHGDTMTVMDPKQ 447
Db 421 SLDDLKYYFFHHSHGDTMTVMDPSR 447

RESULT 4
Q9WVJ3 PRELIMINARY; PRT; 470 AA.
ID Q9WVJ3;
AC Q9WVJ3;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Amino-peptidase (Similar to plasma glutamate carboxypeptidase).
GN PGCP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Liu C.H., Lin B.Y., Chang L.Y.;
RT "Cloning of the mouse aminopeptidase gene.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strassberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body, and Mesonephros;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AF107835; AA043215.1; -
DR EMBL; BC037067; AA037067.1; -
DR EMBL; AK032972; BAC28105.1; -
DR EMBL; AK075686; BAC35891.1; -
DR MEROPS; M28.014; -.
DR MGD; MGI:1889205; PgcP.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR001337; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS00840; PA; 1.
KM Carboxypeptidase.
SQ SEQUENCE 470 AA; 51813 MW; 0F3490681691866A CRC64;

Query Match 88.2%; Score 2133; DB 11; Length 470;
Best Local Similarity 87.7%; Pred. No. 1.8e-154;
Matches 414; Conservative 28; Mismatches 28; Indels 2; Gaps 1;
Qy 1 MKFLIFAFGGVHLISLCSGKAICNGISKRTEFEIEKEIASCGDVAKAIINLAVYGAQ 60
Db 1 MRSFLFFLFI--VHLALSGKAVFRKNGVSQRTFEIKEIEIANVEDVAKAIINLAVYGRQ 58
Qy 61 NRSYERLALVDVTPGPRISGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESA 120
Db 59 NRSYERLGLVDVTPGPRISGSKNLEKAIQIMYQNLQODGLENVHLEQVRIPIHMERGESA 118

```

```

Qy 121 VMLPRILHKLAIILGLSSIGTPPEGITAEVLVWTSPELORRASEARGKITVYNQPYNY 180
Db 119 VMLPRILHKLAIILGLSSIGTPPEGITAEVLVWTSPELORRASEARGKITVYNQPTGX 178
Qy 181 SRTVQYRTQGAVEAKVAGALASLRSVASFISYSPHTGIOYODGVKPIPTACTIVDAE 240
Db 179 EKTQYVAVQGAVEAKVAGALASLRSVASFISYSPHTGIOYODGVKPIPTACTIVDAE 238
Qy 241 MMSHAGHGIKIYVQLKMGAKTYPDTSFNTVAETISGSKYPEQVVLVSGHSDMDVCGA 300
Db 239 MMSHAGHGIKIYVQLKMGAKTYPDTSFNTVAETISGSKYPEQVVLVSGHSDMDVCGA 298
Qy 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEEGGVGAFQYQLHVNINSNLSV 360
Db 299 LDDGGAFISWEALSLIKDGLRPRKTRLVMTAEEGGVGAFQYQLHVNINSNLSV 358
Qy 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLLOPLNTQVLSHGEGDINFWIOAGVGA 420
Db 359 MEADSGTFLPTGLQFTGSDKARAIMEEVMSLLOPLNTQVLSHGEGDINFWIOAGVGA 418
Qy 421 SLDDLKYYFFHHSHGDTMTVMDPKQNTAAAVAVSYVADMEMPLRS 472
Db 419 SLRDLKYYFFHHSHGDTMTVMDPKQNTAAAVAVSYVADMEMPLRS 470

```

```

RESULT 5
Q9ULV0 PRELIMINARY; PRT; 472 AA.
ID Q9ULV0;
AC Q9ULV0;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Liver annexin-like protein.
GN LAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Servillo G., Della Fazio M.A., Piobbico D., Bartoli D., Castellani M.,
RA Brancorsini S., Viola Magni M.;
RT "Lal, a novel gene involved during liver regeneration.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131077; AAF36518.1; -
DR MEROPS; M28.014; -.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR001337; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS00840; PA; 1.
SQ SEQUENCE 472 AA; 52011 MW; 55C68FB90E63F265 CRC64;

```

```

Query Match 87.6%; Score 2117; DB 11; Length 472;
Best Local Similarity 87.5%; Pred. No. 3e-153;
Matches 413; Conservative 24; Mismatches 35; Indels 0; Gaps 0;
Qy 1 MKFLIFAFGGVHLISLCSGKAICNGISKRTEFEIEKEIASCGDVAKAIINLAVYGAQ 60
Db 1 MRSFLFFLFI--VHLALSGKAVFRKNGVSQRTFEIKEIEIANVEDVAKAIINLAVYGRQ 60
Qy 61 NRSYERLALVDVTPGPRISGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESA 120
Db 61 NRSYERLGLVDVTPGPRISGSKNLEKAIQIMYQNLQODGLENVHLEQVRIPIHMERGESA 120
Qy 121 VMLPRILHKLAIILGLSSIGTPPEGITAEVLVWTSPELORRASEARGKITVYNQPYNY 180
Db 121 VMLPRILHKLAIILGLSSIGTPPEGITAEVLVWTSPELORRASEARGKITVYNQPTDX 180
Qy 181 SRTVQYRTQGAVEAKVAGALASLRSVASFISYSPHTGIOYODGVKPIPTACTIVDAE 240
Db 181 GKTQYVBERGAVEAKVAGALASLRSVASFISYSPHTGIOYODGVKPIPTACTITIDAE 240

```

```

Qy 241 MMSHMAHGIXIVIQIKMGAKTYPTDSEFNTVAETSGKYPEQVVLVSGHLSMDVGGGA 300
Db 241 MMSHMAHGIXIVIQIKMGAKTYPTDSEFNTVAETSGKYPEQVVLVSGHLSMDVGGGA 300
Qy 301 MDGGGAFISWEALSLIKDLGRLPRKTLRLVMTAEQGGVGAFOYQOLAKVNISNYSLV 360
Db 301 LDDGGGAFISWEALSLIKDLGRLPRKTLRLVMTAEQGGVGAFOYQOLAKVNISNYSLV 360
Qy 361 MESDAGTFLPTGLQFTSGSEKARAIMBEVMSLLQPLNITQVLSHGSGTDINFMIQAGVPGA 420
Db 361 MEADSGTFLPTGLQFTSGSEKARAIMBEVMSLLQPLNITQVLSHGSGTDINFMIQAGVPGA 420
Qy 421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVVAVVADMEMLPRS 472
Db 421 SLRDDLYKYFFPHSHGDTMTAMDPKQNVAAVAVVAVVADMEMLPRS 472

```

RESULT 6

```

ID 0921Y1 PRELIMINARY; PRT; 472 AA.
AC 0921Y1
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hematopoietic lineage switch 2 related protein.
GN HUS2-RP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Flisher;
RA Chen Y., Talmage D.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF097723; AAC72384.1; -
DR MEROPS; M28.014; -
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR003137; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS50840; PA; 1.
SQ SEQUENCE 472 AA; 51969 MW; DC2111651879662 CRC64;

```

Query Match 87.5%; Score 2116; DB 11; Length 472;
 Best Local Similarity 87.5%; Pred. No. 3.6e-153;
 Matches 413; Conservative 24; Mismatches 35; Indels 0; Gaps 0;

```

Qy 1 MKFLIFAFGGVHLISLCSGKAIKNGISKRTFEEIKEEIASCGDVAKAIINLAVYGRQA 60
Db 1 MRFLFFLVAVVHLPSLGSKAIYKSGVSGRTFGEIKKEIANVEDVAKAIINLAVYGRQA 60
Qy 61 NRSYERLALVDYTPGRLSGSKNLEKAIQIMYQNLQDGLGVLEKVLPRIPMERGESGA 120
Db 61 NRSYERLALVDYTPGRLSGSKNLEKAIQIMYQNLQDGLGVLEKVLPRIPMERGESGA 120
Qy 121 VMLSPRIHKAIIIGLSSIGTPPEGITAEVLVTSFDELQRRASERAGKIIVVNOPTDY 180
Db 121 VMLSPRIHKAIIIGLSSIGTPPEGITAEVLVTSFDELQRRASERAGKIIVVNOPTDY 180
Qy 181 SRTVQRTQGAVEAKVAGALASLIRSVAFSIYSPHGTIOEYQDGVPIPTACTIVEDAE 240
Db 181 SRTVQRTQGAVEAKVAGALASLIRSVAFSIYSPHGTIOEYQDGVPIPTACTIVEDAE 240
Qy 241 MMSHMAHGIXIVIQIKMGAKTYPTDSEFNTVAETSGKYPEQVVLVSGHLSMDVGGGA 300
Db 241 MMSHMAHGIXIVIQIKMGAKTYPTDSEFNTVAETSGKYPEQVVLVSGHLSMDVGGGA 300
Qy 301 MDGGGAFISWEALSLIKDLGRLPRKTLRLVMTAEQGGVGAFOYQOLAKVNISNYSLV 360
Db 301 MDGGGAFISWEALSLIKDLGRLPRKTLRLVMTAEQGGVGAFOYQOLAKVNISNYSLV 360
Qy 361 MESDAGTFLPTGLQFTSGSEKARAIMBEVMSLLQPLNITQVLSHGSGTDINFMIQAGVPGA 420
Db 361 MESDAGTFLPTGLQFTSGSEKARAIMBEVMSLLQPLNITQVLSHGSGTDINFMIQAGVPGA 420

```

```

Db 361 MEADSGTFLPTGLQFTSGSEKARAIMBEVMSLLQPLNITQVLSHGSGTDINFMIQAGVPGA 420
Qy 421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVVAVVADMEMLPRS 472
Db 421 SLRDDLYKYFFPHSHGDTMTAMDPKQNVAAVAVVAVVADMEMLPRS 472

```

RESULT 7

```

ID 070216 PRELIMINARY; PRT; 433 AA.
AC 070216
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hematopoietic lineage switch 2.
GN PCGP OR HUS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams J.H., Chan C.-Y., Klinken S.P.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF009513; AAC17945.1; -
DR MEROPS; M28.014; -
DR MGD; MGI:1889205; PgcP.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR003137; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS50840; PA; 1.
SQ SEQUENCE 433 AA; 47858 MW; 89C9B93712004669 CRC64;

```

Query Match 78.9%; Score 1907; DB 11; Length 433;
 Best Local Similarity 84.8%; Pred. No. 2.9e-137;
 Matches 374; Conservative 25; Mismatches 28; Indels 14; Gaps 2;

```

Qy 1 MKFLIFAFGGVHLISLCSGKAIKNGISKRTFEEIKEEIASCGDVAKAIINLAVYGRQA 60
Db 1 MRFLFFLVAVVHLPSLGSKAIYKSGVSGRTFGEIKKEIANVEDVAKAIINLAVYGRQA 60
Qy 61 NRSYERLALVDYTPGRLSGSKNLEKAIQIMYQNLQDGLGVLEKVLPRIPMERGESGA 120
Db 61 NRSYERLALVDYTPGRLSGSKNLEKAIQIMYQNLQDGLGVLEKVLPRIPMERGESGA 120
Qy 121 VMLSPRIHKAIIIGLSSIGTPPEGITAEVLVTSFDELQRRASERAGKIIVVNOPTDY 180
Db 121 VMLSPRIHKAIIIGLSSIGTPPEGITAEVLVTSFDELQRRASERAGKIIVVNOPTDY 180
Qy 181 SRTVQRTQGAVEAKVAGALASLIRSVAFSIYSPHGTIOEYQDGVPIPTACTIVEDAE 240
Db 181 SRTVQRTQGAVEAKVAGALASLIRSVAFSIYSPHGTIOEYQDGVPIPTACTIVEDAE 240
Qy 241 MMSHMAHGIXIVIQIKMGAKTYPTDSEFNTVAETSGKYPEQVVLVSGHLSMDVGGGA 300
Db 241 MMSHMAHGIXIVIQIKMGAKTYPTDSEFNTVAETSGKYPEQVVLVSGHLSMDVGGGA 300
Qy 301 MDGGGAFISWEALSLIKDLGRLPRKTLRLVMTAEQGGVGAFOYQOLAKVNISNYSLV 360
Db 301 MDGGGAFISWEALSLIKDLGRLPRKTLRLVMTAEQGGVGAFOYQOLAKVNISNYSLV 360
Qy 361 MESDAGTFLPTGLQFTSGSEKARAIMBEVMSLLQPLNITQVLSHGSGTDINFMIQAGVPGA 420
Db 361 MESDAGTFLPTGLQFTSGSEKARAIMBEVMSLLQPLNITQVLSHGSGTDINFMIQAGVPGA 420

```

RESULT 8

076552 ID 076552 PRELIMINARY; PRT; 493 AA.
 AC 076552;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Aminopectidase ES-62 precursor.
 OS Acanthocheilomena viteae (Dipetalonema viteae).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Splirurida; Filarioidae;
 OC Onchocercidae; Acanthocheilomena.
 NCBI_TaxID=6277;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20055923; PubMed=10589978;
 RA Harnett M., Houston K.M., Tate E., Garate T., Apfel H., Adam R.,
 RA Hellem S.M., Panico M., Paxton T., Dell A., Morris H., Brzecki H.,
 RT "Molecular cloning and demonstration of an aminopeptidase activity in
 RT a filarial nematode glycoprotein."
 RL Mol. Biochem. Parasitol. 104:11-23(1999).
 DR EMBL; AF077194; AAC28365.1; -.
 DR MEROPS; M28.015; -.
 DR InterPro; IPR003137; PA.
 DR PROSITE; PS50840; PA; 1.
 KW Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 493 AMINOPEPTIDASE ES-62.
 SQ SEQUENCE 493 AA; 54916 MW; C7E9227DD18E1DF4 CRC64;
 Query Match 31.4%; Score 759.5; DB 5; Length 493;
 Best Local Similarity 36.6%; Pred. No. 1.7e-49;
 Matches 174; Conservative 92; Mismatches 166; Indels 43; Gaps 13;
 QY 15 LSLGSGKAIC-----KNGSKRTFEIEKEINASCGBVAKAIINLVAGQKRSYERL 67
 DB 14 LTVVLGAVALPDKTVPAPKNYIQETFGKVAE-----LQYITKGEEVGLAYQWL 62
 QY 68 ALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPY-RIPMERGESAVMLEPR 126
 DB 63 SKLVDPGGRHNVGSDLSKSIAPLEBSIKNDNPKMHEEPNLPHWKRGNDVEMIEPR 122
 QY 127 IHKTAIIGLSSISITPREGITAEVLVYTSFDELQRRASEARKIVVNOPIYNSRTVQY 186
 DB 123 NQRLNVLAIG---GSEPASATGEVTVIYDLDDV--KPDVDVGIKIVTQATGAGYPLTLKY 177
 QY 187 RTQGAVERAKYKALASLIRSVASFSIYSPHNGIOEQGVCKIPACTIVDAMSGMA 246
 DB 178 R-RSVKLFEOJGALGVVKSITSPINSPTGTGAEN--TTTPACLTIEAEMLERLY 233
 QY 247 SHGIRIVQLKMGAKTYPDTSFNTVAEITGSKYPEQVVLVSGHLDSDVQAGAMDGGG 306
 DB 234 RSGKGIYIRMDKSHYEPINSNLIFITGSEKPSVVLVSAVDSMDVQAGALDDGAG 293
 QY 307 AFISWEALSLIKDGLG---LRPKRTLVLTAEBOGGVAFQYQHLKYNL-SNYSLV 360
 DB 294 CAVVMSALHSIKLAERPKFKPKRTIGIFMTSEBOGYGAKHYIYHKNDSEPKFYFV 353
 QY 361 MESDAGTLPFG---LQFTGSEKARAMEEVMSLQDLNITQVL--SHSGGTINFPIQ 414
 DB 354 SETPTGFKSTNMLAHSFSGDKKSMLEKIEITRLSRNGALGLINSVVG-DVTFFAK 412
 QY 415 AGVPGASLDD--LYKYFFPHSHSDTWTVMDPKQNVAAVMAVSVVAVDMEB 467
 DB 413 DGISVNTIPKAVDYIYFFHHTAGDVTVLKQDLETTTISFATLGAVIANMD 467

RESULT 9
 ID 09UNM8 PRELIMINARY; PRT; 140 AA.
 AC 09UNM8;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Aminopeptidase (Fragment).

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Liu C.H., Lin B.Y., Chang L.Y.;
 RT "Cloning of the human aminopeptidase gene."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF107833; AAD43213.1; -.
 DR MEROPS; M28.014; -.
 DR InterPro; IPR003137; PA.
 DR PROSITE; PS50840; PA; 1.
 FT NON_TER 1 1
 FT NON_TER 140 140
 SQ SEQUENCE 140 AA; 15472 MW; F34ACEAB33A1AD24 CRC64;
 Query Match 29.0%; Score 701; DB 4; Length 140;
 Best Local Similarity 100.0%; Pred. No. 6.8e-46;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 78 LSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIPMERGESAVMLEPRIHKTAIIGLGS 137
 DB 1 LSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIPMERGESAVMLEPRIHKTAIIGLGS 60
 QY 138 SIGPREGITAEVLVYTSFDELQRRASEARKIVVNOPIYNSRTVQYTGAVEAKY 197
 DB 61 SIGPREGITAEVLVYTSFDELQRRASEARKIVVNOPIYNSRTVQYTGAVEAKY 120
 QY 198 GALASLIRSVASFSIYSPHT 217
 DB 121 GALASLIRSVASFSIYSPHT 140

RESULT 10
 ID 08EBH7 PRELIMINARY; PRT; 414 AA.
 AC 08EBH7;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Peptidase, M28D family.
 GN SO3539.
 OS Shewanella oneidensis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
 OC Alteromonadaceae; Shewanella.
 NCBI_TaxID=70863;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MR-1;
 RX MEDLINE=22297686; PubMed=12368813;
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
 RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
 RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
 RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
 RA Mueller T.V., Smith H.O., Venter J.C., Nealeon K.H., Fraser C.M.;
 RA "Genome sequence of the dissimilatory metal ion-reducing bacterium
 RT Shewanella oneidensis";
 RL Nat. Biotechnol. 20:1118-1123(2002).
 DR EMBL; AE015790; AAN56530.1; -.
 DR TIGR; SO3539; -.
 KW Complete proteome.
 SQ SEQUENCE 414 AA; 44028 MW; C9AEB2A5773DBC84 CRC64;
 Query Match 27.1%; Score 655; DB 16; Length 414;
 Best Local Similarity 36.4%; Pred. No. 1.3e-41;
 Matches 148; Conservative 72; Mismatches 173; Indels 14; Gaps 4;
 QY 74 VGRPLSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIPMERGESAVMLEPRIHKTAIIL 133

```

Db      3 VGPRLGSPKQVIAVNMAMNKLTSIGFDKVKYEPQVPIWMEGKAKIISPVEQPLVIT 62
Qy      134 GLGSSIGTPREGITAEVLVTSFDELQRR--SEANGKIVTVNQPPI-----NYRRTQY 186
Db      63 ALGGSVAAPVGVIGKIKARFNSLEALQATPDVDVGLAIFIDQKTRHITGEGYKSVGG 122
Qy      187 RTQGAWEAKVAGALASLIRSVASFISYSPHTGISOYQGVPKIPACTIVEDAEWMSMA 246
Db      123 RSKGAIIAAOKGAVAVITIRISGTDHDMAHGIMHYODGVPKIPAAANSNPADLVDM 182
Qy      247 SHGKIVIQLMKGAKTYPDTSFNTVAETGSKYPEQVVLVSGHLSMDVGGAMDGGG 306
Db      183 KRDPNAVLELNNSPKDLGNTSYNVIAEVTGSSKNEIYLGAHLDSMDEGTGALIDCGG 242
Qy      307 AFISHEALSLIKDLRPRKTRILVMTAEQGVGAQOYOLHKVNISNYSLVMSDAG 366
Db      243 IAIYAAAANKHIDLPQKAPKRTIRVLYAAEEMGLIGKAYAEAKHTELPLHYIAESDFG 302
Qy      367 TELPTGLQFTGSEKARAIMEEVMSLLOPLNITQVL-----SHGEGTDINFWIQAGVPGASL 422
Db      303 AGPIYQITTKVNEK--VPAQVQESIKPMYNGVALGDNQASGGPDISMPLALGVVASL 359
Qy      423 LDDLKYKFFHSHGDTMTVMDPKQMNVAAMVAVSYVADMEEM 469
Db      360 RODGRDYDFYHHTPNDTLDKINPKALQAVAAVAFAYIMANSIVL 406

```

RESULT 11

```

Q9A5A8      PRELIMINARY;      PRT;      467 AA.
AC      Q9A5A8;
DT      01-JUN-2001 (TREMBLrel. 17, Created)
DT      01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Amino peptidase, putative.
GN      CC2544.
OS      Caulobacter crescentus.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC      Caulobacteraceae; Caulobacter.
OX      NCBI_TaxID=155892;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 19089 / CB15;
RX      MEDLINE=21173698; PubMed=11259647;
RA      Nieman W.C., Fejdllyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA      Eiesen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA      Potocka I., Nelson W.C., Newton A., Stephens C., Padke N.D., Ely B.,
RA      DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA      Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA      Ueberbach T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA      Salzberg S.L., Venter J.C., Shapiro L., Frazer C.M.;
RT      "Complete genome sequence of Caulobacter crescentus."
RL      Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR      EMBL; AE005922; AAK24515.1; -.
DR      TIGR; CC2544; -.
DR      InterPro; IPR003137; PA.
DR      PROSITE; PS00840; PA. 1.
KW      Complete proteome.
SQ      SEQUENCE 467 AA; 48918 MW; 9984DE2A9D84F00F CRC64;

```

Query Match 26.5%; Score 640; DB 16; Length 467;
 Best Local Similarity 36.2%, Pred. No. 2.2e-40;
 Matches 147; Conservative 72; Mismatches 177; Indels 10; Gaps 4;

```

Qy      70 LVDYGPRLSGSKNLEKAIQIMYQNLQODGEKVALEPVRLPHMERGESAVMLEPRIMK 129
Db      49 LTTNIGPRLVSGPAAKAKKQMSVAKFALGFNINIVDEAKRSWREGSSALVAPYMK 108
Qy      130 IAILGLSSISGTPREGITAEVLVTSFDELQRRASEA-RGKIVTVNQPPIYNSRTVQY-- 186
Db      109 LGAVGLGRVSTPAGIGIEAEVALFTFTADMAAPDGAALGKIKIVITQPVWRTONGAGYGA 168

```

```

Qy      187 ----RTQGAWEAKVAGALASLIRSVASFISYSPHTGISOYQGVPKIPACTIVEDAEWM 242
Db      169 AGVSRRIRGPVPAARAGVAMVILIRISITSDSVPHPTGTGASEGVVTPAALIGVPAEQ 228
Qy      243 SRMASHGKIVIQLMKGAKTYPDTSFNTVAETGSKYPEQVVLVSGHLSMDVGGAMD 302
Db      229 ERLAARKVPMKVKLTSSVPMNVANNISGDDIGSKSPDEVIVIGHLSMDVGTGALD 288
Qy      303 DGGGAFISWEALSLIKDLRPRKTRILVMTAEQGVGAFOYOLHKVNISNYSLVME 362
Db      289 DATGIALTAAKLIQDLPKAPKRTIRVMMGSESGG--SEAYLAANKALSTMVLAGE 347
Qy      363 SDAGTFPLTGLQFTGSEKARAIMEEVMSLLOPLNI--TVLSHGEGTDINFWIQAGVPGA 420
Db      348 SDTGADNIIYSLQVPAGSLDHHVVAASVLAFLKTIYDRPFAAGADVSGIERAGVPVI 407
Qy      421 SLDDLKYKFFHSHGDTMTVMDPKQMNVAAMVAVSYVADME 466
Db      408 NLNQDASRFYDHTMDTDLKVPAPAEIAQVAAWTSILVLYLVADSD 453

```

RESULT 12

```

Q9R114      PRELIMINARY;      PRT;      139 AA.
AC      Q9R114;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Amino peptidase (fragment).
GN      PGCP.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Liver;
RA      Liu C.H., Lin B.Y., Chang L.Y.;
RT      "Cloning of the mouse aminopeptidase gene."
RL      Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF107832; AAD3212.1; -.
DR      HSSP; P80561; 1X00.
DR      MEROPS; M28.014; -.
DR      MGD; MGI:1889205; Pgcp.
FT      NON_TER 1
FT      NON_TER 139
SQ      SEQUENCE 139 AA; 15161 MW; 3B21B7CF2D2DE648 CRC64;

```

Query Match 25.8%; Score 623; DB 11; Length 139;
 Best Local Similarity 84.7%, Pred. No. 6.2e-40;
 Matches 116; Conservative 14; Mismatches 7; Indels 0; Gaps 0;

```

Qy      277 GSKYPEQVVLVSGHLSMDVGGAMDGGAFISWEALSLIKDLRPRKTRILVMTAE 336
Db      3 GSMYPEEVVLVSGHLSMDVGGALDDGGGFISWEALSLIKDLRPRKTRILVMTAE 62
Qy      337 EGGGVGAFOYOLHKVNISNYSLVMSDAGTFPLTGLQFTGSEKARAIMEEVMSLLOPLN 396
Db      63 EGGIGASQYVELHKANISKSLVMEADSGTFPLTGLQFTGSDKARAIMEEVMSLLOPLN 122
Qy      397 ITQVLSHGEGTDINFWI 413
Db      123 VTKVFSNDEGTDISFRI 139

```

RESULT 13

```

Q9PHE8      PRELIMINARY;      PRT;      472 AA.
AC      Q9PHE8;
DT      01-OCT-2002 (TREMBLrel. 22, Created)
DT      01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Aminopeptidase.

```


GN XAC3309.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 CX NCBI_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavan F., Cardoso J., Chamberggo F., Clapina L.P.,
 RA Ciarrelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truifi D., Tsai S.M., White F.F.,
 RA "Seubel J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL: AEO11976; AAA38152.1;
 DR InterPro: IPR00137; PA.
 DR PROSITE: PS50840; PA; 1.
 KW Complete proteome.
 SQ SEQUENCE 472 AA; 48788 MW; 9956CDF1F49FC6B8 CRC64;

 Query Match 20.7%; Score 499.5; DB 16; Length 472;
 Best Local Similarity 29.7%; Pred. No. 1.2e-29;
 Matches 140; Conservative 76; Mismatches 218; Indels 37; Gaps 10;

 QY 12 VHLTLSCGKAIKNGI---SKRTFEIKERIASCGVAKAIIIMAVYKQNSYERLA 68
 DB 8 ISALLASCALAAATSIIPDSALRTAQLRE-ALADDTGFAVVO----- 50
 QY 69 LLDVTGPRLSGSKNLEKAIQIMYQNLQDGLKRVHLEPVRIIPMERGESAVMLEPRRIH 128
 DB 51 SLTEVGPRLAGBADPRAVAMAKAKFASLGFDKVMTEPVFPKMERSEHAIVGAAHQ 110
 QY 129 KIALIGSSIGTPREGITAEVLVVTSPFDELQRR-ASEARKIVVYNQPYI-----NYS 181
 DB 111 PLITLALGSGPGGTVEG---EVRFEETLAAIQAPASGLAKIAFVYQWTKARDGXDYG 167
 QY 182 RTVOYRTQGAVEAKVAGLASLRSVASFISYSPHTGIGQYQDGVPKITPACTIVEDAEM 241
 DB 168 NGAVNRKSGSEARIRKGAIGFVMSAGTDSHRVPHGTITRFEDELTPVPAALVDPDNO 227
 QY 242 MSRWASHGIKIVIOIKMGAKTYPTDTSFNTVAIRITGSKYRPOVVLVSGHLSMDVGOGAM 301
 DB 228 LARITALG-SARVRLALDCGMDGATSYNVIGETIGRSKREVVVIGHLSMDLGTGAI 286
 QY 302 DDGGATISWEALSLIKDLGIRPKRTLRLVMTABEGCGVAFQYQLH---KYNISNYS 358
 DB 287 DDGAGVAITMAAGHLIQQLKQPRKRTIRVAFANBEOGLHGKAYAAHGDADMDALHQ 346
 QY 359 LVMSRD--AGFTPLTGLOFTGSEKARAIMEEVMSLLOPLINTQYLSH-GESTDINPMTQA 415
 DB 347 IGASDPCGAGRIVAFNFGAALPDDSRATKQIAVLAPLGIAYEPSSKGGPPDVGPIASAK 406
 QY 416 GVPASLLDLKYYFFPHSHSGDTMTVMDPKQNVAAVAWAVSVVADME 466
 DB 407 GGAGWGLAQDGTGYFDLHTADDTLKDIPKALQNVAAAYVFAVLAEND 457

AC Q8P625;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Xanthopeptidase.
 GN XCC3157.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 CX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavan F., Cardoso J., Chamberggo F., Clapina L.P.,
 RA Ciarrelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truifi D., Tsai S.M., White F.F.,
 RA "Seubel J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL: AEO12431; AAA42427.1;
 DR InterPro: IPR00137; PA.
 DR PROSITE: PS50840; PA; 1.
 KW Complete proteome.
 SQ SEQUENCE 375 AA; 38756 MW; C75FA4A914E2EA05 CRC64;

 Query Match 17.4%; Score 420; DB 16; Length 375;
 Best Local Similarity 33.0%; Pred. No. 9.9e-24;
 Matches 105; Conservative 54; Mismatches 145; Indels 14; Gaps 5;

 QY 59 AONRSYERLALVDVTGPRLSGSKNLEKAIQIMYQNLQDGLKRVHLEPVRIIPMERGEE 118
 DB 41 ADSTGFAYVSLTTEVGPRLAGBADPRAVAMAKAKFASLGFDKVMTEPVFPKMERSE 100
 QY 119 SAVMLEPRIRKIALIGSSIGTPREGITAEVLVVTSPFDELQRR-ASEARKIVVYNQPY 177
 DB 101 OAAVIGAAQPLHITLALGSGPGGTVEG---EIVRFEETLAAIQAPASGLAKIAFVYQW 157
 QY 178 I-----NYSRTVOYRTQGAVEAKVAGLASLRSVASFISYSPHTGIGQYQDGVPKIPT 231
 DB 158 VKARDGDYNGAVNRKSGSEARIRKGAIGFVMSAGTDSHRVPHGTITRFEDELTPVPA 217
 QY 232 ACTIVEDAEMMSRWASHGIKIVIOIKMGAKTYPTDTSFNTVAIRITGSKYRPOVVLVSGH 291
 DB 218 AALSVPDANOALARIALG-STVRRLALDCGMDGATSYNVIGETIGRSKREVVVIGH 276
 QY 292 DSDMDVGAGADDGGAFTISWEALSLIKDLGIRPKRTLRLVMTABEGCGVAFQYQLH 351
 DB 277 DSDMDLGTGALIDDGAGVAITMAAGHLIQQLKQPRKRTIRVAFANBEOGLHGKAYAAH 336
 QY 352 VNISNYSLVN---ESDAG 366
 DB 337 KDAKMDMALHQIGASDFG 354

RESULT 14
 Q8P625
 ID Q8P625
 PRELIMINARY; PRT; 375 AA.

RESULT 15
 Q93EJ5
 ID Q93EJ5
 AC Q93EJ5
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:49:10 ; Search time 37.9351 Seconds
(without alignments)
526.445 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417
Sequence: 1 MKFLIFAFGCVHLLSLCSG.....AVNAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A COMB .pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B COMB .pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A COMB .pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B COMB .pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS COMB .pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	100.0	473	US-09-482-273-120	Sequence 120, App
2	229.5	9.5	496	US-09-079-955-2	Sequence 2, Appli
3	226.5	9.4	481	US-09-330-095-1	Sequence 1, Appli
4	182	7.5	584	US-09-252-991A-18292	Sequence 18292, A
5	165	6.8	33	US-09-482-273-224	Sequence 224, App
6	165	6.8	34	US-09-482-273-187	Sequence 187, App
7	165	6.8	750	US-08-325-553-2	Sequence 2, Appli
8	165	6.8	750	US-08-394-152A-2	Sequence 2, Appli
9	165	6.8	750	US-09-044-668-2	Sequence 2, Appli
10	165	6.8	750	US-08-705-477E-2	Sequence 2, Appli
11	161	6.7	693	US-08-705-477E-101	Sequence 101, App
12	160	6.6	707	US-08-528-122-18	Sequence 18, Appl
13	160	6.6	707	PCT-US95-11720-18	Sequence 18, Appl
14	143.5	5.9	622	US-08-547-197-1	Sequence 1, Appli
15	143.5	5.9	622	US-08-957-940-1	Sequence 1, Appli
16	137.5	5.7	537	US-09-079-955-5	Sequence 5, Appli
17	126	5.2	760	US-08-547-197-2	Sequence 2, Appli
18	126	5.2	760	US-08-957-940-2	Sequence 2, Appli
19	108.5	4.5	439	US-09-634-238-412	Sequence 412, App
20	102.5	4.2	337	US-09-107-532A-5998	Sequence 5998, Ap
21	100.5	4.2	647	US-09-423-439-60	Sequence 60, Appl
22	100.5	4.2	4928	US-09-036-987A-5	Sequence 5, Appli
23	100.5	4.2	4928	US-09-370-700-5	Sequence 5, Appli
24	100.5	4.2	4928	US-09-603-207-5	Sequence 5, Appli
25	100	4.1	2325	US-08-417-089-6	Sequence 6, Appli
26	100	4.1	2325	US-08-695-651-6	Sequence 6, Appli
27	100	4.1	2325	US-08-930-285-6	Sequence 6, Appli

28	100	4.1	2325	US-08-695-421-6	Sequence 6, Appli
29	100	4.1	2325	US-08-697-826A-10	Sequence 10, Appl
30	98.5	4.1	934	US-09-252-991A-26779	Sequence 26779, A
31	97.5	4.0	403	US-08-776-246-4	Sequence 4, Appli
32	97.5	4.0	619	US-09-107-532A-4897	Sequence 4897, Ap
33	96	4.0	331	US-09-724-623-78	Sequence 78, Appl
34	96	4.0	465	US-09-634-238-411	Sequence 411, App
35	96	4.0	693	US-08-463-620-11	Sequence 11, Appl
36	96	4.0	693	US-08-124-917-11	Sequence 11, Appl
37	96	4.0	693	US-08-914-853-11	Sequence 11, Appl
38	96	4.0	693	PCT-US95-03934A-11	Sequence 11, Appl
39	95.5	4.0	340	US-09-789-300A-2	Sequence 2, Appli
40	95.5	4.0	373	US-09-198-452A-416	Sequence 416, App
41	95.5	4.0	406	US-09-134-001C-3544	Sequence 3544, Ap
42	94.5	3.9	1452	US-09-252-991A-18118	Sequence 18118, A
43	94	3.9	455	US-09-328-352-5839	Sequence 5839, Ap
44	93	3.8	450	US-09-107-532A-6534	Sequence 6534, Ap
45	92.5	3.8	440	US-09-107-532A-4071	Sequence 4071, Ap

ALIGNMENTS

RESULT 1
US-09-482-273-120
Sequence 120, Application US/09482273
Patent No. 6534631
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
EARLIER FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
NUMBER OF SEQ ID NOS: 267
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 120
LENGTH: 473
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: SITE
LOCATION: (473)
OTHER INFORMATION: Xaa equals stop translation
US-09-482-273-120
Query Match 100.0%; Score 2417; DB 4; Length 473;
Best Local Similarity 100.0%; Pred. No. 1.6e-232;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKFLIFAFGCVHLLSLCSGKAICKNGISKRTEEIKKEIASCGDAKAIINLAIVYKAO 60
Db 1 MKFLIFAFGCVHLLSLCSGKAICKNGISKRTEEIKKEIASCGDAKAIINLAIVYKAO 60
QY 61 NRSYERIALVDPVGPRLSSKNLEKAIQIMYONLQODGLEKHLPEVRIPHERGESA 120
Db 61 NRSYERIALVDPVGPRLSSKNLEKAIQIMYONLQODGLEKHLPEVRIPHERGESA 120
QY 121 VMLEPRHKAIAIGLSSISGTPPEGTAIVLVVTSFDELQRRASEARGKIVVNOPIYNY 180
Db 121 VMLEPRHKAIAIGLSSISGTPPEGTAIVLVVTSFDELQRRASEARGKIVVNOPIYNY 180
QY 181 SRTVQVTRGAVDAARVAGALASLIRSVASFISYSPHTGIEYODGVKIPYACTIVDAE 240
Db 181 SRTVQVTRGAVDAARVAGALASLIRSVASFISYSPHTGIEYODGVKIPYACTIVDAE 240

Query 241 MMSRASHGKIKVIQIKKGAATYPTDTSFNTVAETGSKYEQVVLVSGHLSMDVGGCA 300
Db 241 MMSRASHGKIKVIQIKKGAATYPTDTSFNTVAETGSKYEQVVLVSGHLSMDVGGCA 300
Qy 301 MDDGGAFISWEALSLIDGLRPRTTLRLVMTAEQGGVGAFOYQOLHKVNISNYSLV 360
Db 301 MDDGGAFISWEALSLIDGLRPRTTLRLVMTAEQGGVGAFOYQOLHKVNISNYSLV 360
Qy 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLLQPLNTQVLSHGEGTINFWIQGVGA 420
Db 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLLQPLNTQVLSHGEGTINFWIQGVGA 420
Qy 421 SLDDLKYKFFPHSHGDTMTVMDPKOMVAAVAVSVYVADMEMLPRS 472
Db 421 SLDDLKYKFFPHSHGDTMTVMDPKOMVAAVAVSVYVADMEMLPRS 472

RESULT 2
US-09-079-955-2
Sequence 2, Application US/09079955A
Patent No. 6465209
GENERAL INFORMATION:
APPLICANT: Alexander Blinkovsky
APPLICANT: Kimberly Brown
APPLICANT: Elizabeth Golligly
APPLICANT: Tony Byun
APPLICANT: Thomas Mathiasen
APPLICANT: Lene V. Kotod
APPLICANT: Mikio Fujii
APPLICANT: Chigusa Shizuoka
TITLE OF INVENTION: Methods For Producing Protein
FILE REFERENCE: 5253,500-US
CURRENT APPLICATION NUMBER: US/09/079,955A
CURRENT FILING DATE: 1998-05-15
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 496
TYPE: PRT
ORGANISM: Aspergillus oryzae
US-09-079-955-2

Query Match 9.5%; Score 229.5; DB 4; Length 496;
Best Local Similarity 22.6%; Pred. No. 4,1e-14;
Matches 96; Conservative 83; Mismatches 175; Indels 71; Gaps 19;

Qy 77 RLGSKNLEKAIQIMQNLQDDLEKHLPEVRIPIHMERGESAVMLEPRHIAIILGIG 136
Db 53 RYFGKADHDVTNLYLKEKKTGYDVYKQ-QVHLSMADOTLKVGBEIE-----AK 105
Qy 137 SSIGTPREGITAEVLVVTSPDELQ-RASEARGKIVVYNQ--PYINTSRVYQTQAGV 192
Db 106 TMTYSPSEVETADVAVVKNLGCSEADYPSDVEGKVALIKRGCEPFGDGS-----V 155
Qy 193 EAKGAGLASLIRSVAISFISPHGTIGQYODGVKIPACTIVTDAEMSMASHGKI 252
Db 156 LAARKKAASIVYNNVAGSMAGTLGAQ--SDKGPYSAIVGISLEDDQKLILDAAG-SV 212
Qy 253 VIQTMGAATYPTDTSFNTVAETGSKYEQVVLVSGHLSMDVGGAMDDGGAFISWE 312
Db 213 SVDLVAVDSKQENRT--TYNVVAQTKGSD-PNNVVALGCHTDSVEARGINDSGIISNLV 270
Qy 313 ALSIKLIDGLRPKRTTLRLVMTAEQGGVGAFOY-----QHLKVN-----SN 356
Db 271 IAKALTOYSV--KNAVRFLEMTAEERFGLGS--NYVSHLNATELNKIRLYLNFDMIASPN 327
Qy 357 YSL-VMESDAGTFLPTGLQFTGSEKARAIMEEVMSLLQPLNTQVLSHGEG-TDINFWIQ 414
Db 328 YALMTYDGGSAFNOSGP--AGSAQIEKLFED--YYSIDILPHIPTQPGRSDYEAFIL 382
Qy 415 AGVPGASLIDDL-----YKYFFHSHGDTMTVMDPKQ--MNVAAAVW 455

Db 383 NGIPSGFLTGABGIMSEENASRWGQAGVAYDANYHAAGDMNTLNHAEFLINSKATAF 442
Qy 456 AVVSY 460
Db 443 AVATY 447

RESULT 3
US-09-330-095-1
Sequence 1, Application US/09330095
Patent No. 6127161
GENERAL INFORMATION:
APPLICANT: Kikkoman Corporation
TITLE OF INVENTION: Leucine Aminopeptidase Gene, Recombinant DNA, and
FILE REFERENCE: PH-622
CURRENT APPLICATION NUMBER: US/09/330,095
CURRENT FILING DATE: 1999-06-11
EARLIER APPLICATION NUMBER: JP-164611/1998
EARLIER FILING DATE: 1998-06-12
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 481
TYPE: PRT
ORGANISM: Aspergillus oryzae
US-09-330-095-1

Query Match 9.4%; Score 226.5; DB 3; Length 481;
Best Local Similarity 22.0%; Pred. No. 7.8e-14;
Matches 93; Conservative 84; Mismatches 180; Indels 65; Gaps 18;

Qy 77 RLGSKNLEKAIQIMQNLQDDLEKHLPEVRIPIHMERGESAVMLEPRHIAIILGIG 136
Db 38 RYFGKADHDVTNLYLKEKKTGYDVYKQ-QVHLSMADOTLKVGBEIE-----AK 90
Qy 137 SSIGTPREGITAEVLVVTSPDELQ-RASEARGKIVVYNQPYINTSRVYQTQAGVAA 195
Db 91 TMTYSPSEVETADVAVVKNLGCSEADYPSDVEGKVALIKRGCEAFG-----DKSVLAA 143
Qy 196 KVGAGLASLIRSVAISFISPHGTIGQYODGVKIPACTIVTDAEMSMASHGKIKVIQ 255
Db 144 KAKAASIVYNNVAGSMAGTLGAQ--SDKGPYSAIVGISLEDDQKLILDAAG-SVSD 200
Qy 256 LKMAKTYPTDTSFNTVAETGSKYEQVVLVSGHLSMDVGGAMDDGGAFISWEALS 315
Db 201 LWDVSKQENRT--TYNVVAQTKGSD-PNNVVALGCHTDSVEARGINDSGIISNLVAV 258
Qy 316 LIKDLGRPRTTLRLVMTAEQGGVGAFOY-----QHLKVN-----SNYSL 359
Db 259 ALTOYSV--KNAVRFLEMTAEERFGLGS--NYVSHLNATELNKIRLYLNFDMIASPNVAL 315
Qy 360 -VMESDAGTFLPTGLQFTGSEKARAIMEEVMSLLQPLNTQVLSHGEG-TDINFWIQAGV 417
Db 316 MTDGDSAFNOSGP--AGSAQIEKLFED--YYSIDILPHIPTQPGRSDYEAFILNGI 370
Qy 418 PGASLIDDL-----YKYFFHSHGDTMTVMDPKQ--MNVAAAVW 458
Db 371 PAGGLFGAGGIMSEENASRWGQAGVAYDANYHAAGDMNTLNHAEFLINSKATAFAVA 430
Qy 459 SY 460
Db 431 TY 432

RESULT 4
US-09-252-991A-18292
Sequence 18292, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18292
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18292

```

```

Query Match
Best Local Similarity 7.5%; Score 182; DB 4; Length 584;
Matches 95; Conservative 61; Mismatches 141; Indels 106; Gaps 19;

QY 65 ERLALVDYVGPRLSGSKLEKAIQIMYONLQODGLEKYLEPVRIPIHMERGESAVMLE 124
DB 123 EDIASLND--GNRAAATPGYQASVDYVKOTLQXAGY-KVSVPPEPTAY-----Y 169
QY 125 PRIRIKAILGSSIGTPE-----GTTAVLVV-----TS 155
DB 170 PK-----GPGSLSATVPQPVYEWKDFYLSQTEAGDVTAKVPVDSLGAQNTSTS 222
QY 156 FDELORPASEARGLIVVNOPIYVSRVQYRTQGAVERAAKVALSIRVASFSIYSP 215
DB 223 GCEADPRANFPAGSIALIQRTCNFE-----QKELAAAGAGVITFQGN---TDD 272
QY 216 HTGI-----OEYODGVKPIPTACITVEDAEMSRMASHGKIYIQLMGAKTPDITDSF 269
DB 273 RKGLENTVSGSYGGIPLV---FATYDNGVAMSGTPQLQHLVVDV---VAKIETRY 324
QY 270 NTVAEITGSKRPEQVYVSGHLSNDVQGAANDGCGAFISWE-ALSLIKDLGRLPKRTL 328
DB 325 NVVAE-TRRGNPNVVMVGAHLDSVFEGPGINDGSGAQLLEMAVLAKAL---PVAKV 380
QY 329 RLVLMTAEQGGVGAFOYQLOHKVNISNYSLVMSDATTPLPTGLQFSGSEKARAIMREV 388
DB 381 RFAMWGAEEAGLVGSTRHYQ-----NLAPER-----KKIKIAYLN-- 415
QY 389 MSLLQPLNITQVLSHGEGTDINFWIOAGVPGASLLDLKYKFF 431
DB 416 FDMIGSPFNGFIYDGGSD--FGLQ-GPPGSAIERLFEAYF 455

```

```

RESULT 5
US-09-482-273-224
; Sequence 224, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 224
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-224

```

```

Query Match
Best Local Similarity 6.8%; Score 165; DB 4; Length 33;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 440 MTVMDFKQNVAAVAVSVYVADMEMLPRS 472
DB 1 MTVMDFKQNVAAVAVSVYVADMEMLPRS 33

```

```

RESULT 6
US-09-482-273-187
; Sequence 187, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 187
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals stop translation
US-09-482-273-187

```

```

Query Match
Best Local Similarity 6.8%; Score 165; DB 4; Length 34;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 440 MTVMDFKQNVAAVAVSVYVADMEMLPRS 472
DB 1 MTVMDFKQNVAAVAVSVYVADMEMLPRS 33

```

```

RESULT 7
US-08-325-553-2
; Sequence 2, Application US/08325553
; Patent No. 5538866
; GENERAL INFORMATION:
; APPLICANT: Israeli, Ron S.
; APPLICANT: Heston, Warren D.W.
; APPLICANT: Fair, William R.
; TITLE OF INVENTION: THE PROSTATE-SPECIFIC MEMBRANE ANTIGEN
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,553
; FILING DATE:

```

CLASSIFICATION: 435
 PRIOR APPLICATION DATA: US/07/973,337A
 APPLICATION NUMBER: 05 NOV 1992
 FILING DATE: 05 NOV 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 1747/41426
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 977-9550
 TELEFAX: (212) 664-0525
 TELETYPE: 422523 COOP UI
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 750 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-325-553-2

Query Match 6.8%; Score 165; DB 1; Length 750;
 Best Local Similarity 19.2%; Pred. No. 2.2e-07;
 Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

3 FLIFAFGCVHLLSLCSGKAIKNGISKRTFEEIKELASGDAVKAIINLAVYGKQNR 62
 32 FFLGFLGFWFKSKSNEATNITPRKMKAFLELKAEE-----NIKKFLYNF----- 77
 63 SYERLALLVDIVGPRISGSK--NLEKAIQIOMYONLOODGLEKVLHPV-----RIPHW 113
 78 -----TQIPLHAGTEQNFOLAKQIOSQWKEFGDLSVELAHYDVLSPKNTHPNY 127
 114 -----ERGEE--SAVMLEPRHIAIKAILGLSSIGTPPEGITAEVLVYTSFDELQRRASEA 166
 128 ISIINDEGNEIFNTSLFEP-----PPGYENVSQDIVPPFSAFSPQGM-P 170
 167 RGIIVVNOPYINYSRT-----VOYRTQGAVEAAKVG-----ALASLIRSV 207
 171 EGDLYV-----VYNVARTEDFFKLERDMKINCSGKIVIAHYGVFRGNKVKNAQLAGAGV 225
 208 ASFS-----IYSPHTGIGQYOD----- 224
 226 ILYSDPADYFAP--GVKSYPDGMNLPGGGVORGNILINLNGADPLTPGYANEVAYRRGI 283
 225 -----GVKPIPTACTIVEDAEMMSMAASHGKIVIQKGAATYD----- 265
 284 AEAVGLPEIPVHPPIGYDDAQKLE-----KKGGSAPPDSSWRGSLKVPYNVGP 331
 266 -----TDSFNTVAEITGSKYEQVVLVSGHLSMDVGGAMDG 304
 332 GFTGNFSTQVKVMHISTNEVTRINVIQTLRGAVEBDRYVILGHRDSWVF--GGIDPQ 389
 305 GGAFISWE--ALSLIKDLGLRPKRTLRVLWMTAEBOGVGAFOYQOLHKVNISVSLVM 361
 390 SGAIVVHEIYVSFGTLKKEGWRPRRTILFASWDAEEFGILGSTMAE-----ENSRLLQ 443
 362 ESDAGTFLPTGLQFTGSEKARAIMIEWM-SLLQPLNTIYOVLSHGEGTDINFIQAGVGA 420
 444 ERGV-AVINADSSIEGNVTLRVDCPTLMTSLVH--NLTKELKSPD-----EGFEKG 491
 421 SLDDLYK 428
 492 SLVESWTK 499

RESULT 8
 US-08-394-152A-2
 Sequence 2, Application US/08394152A
 Patent No. 5935818
 GENERAL INFORMATION:
 APPLICANT: Israeli, Ron S.
 APPLICANT: Heston, Warren D.W.
 APPLICANT: Fair, William R.

TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND
 TITLE OF INVENTION: USES THEREOF
 NUMBER OF SEQUENCES: 48
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States of America
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM 330 466 DX2
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/394,152A
 FILING DATE: 24-FEB-95
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 41426-B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 750 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-394-152A-2

Query Match 6.8%; Score 165; DB 2; Length 750;
 Best Local Similarity 19.2%; Pred. No. 2.2e-07;
 Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

3 FLIFAFGCVHLLSLCSGKAIKNGISKRTFEEIKELASGDAVKAIINLAVYGKQNR 62
 32 FFLGFLGFWFKSKSNEATNITPRKMKAFLELKAEE-----NIKKFLYNF----- 77
 63 SYERLALLVDIVGPRISGSK--NLEKAIQIOMYONLOODGLEKVLHPV-----RIPHW 113
 78 -----TQIHLHAGTEQNFOLAKQIOSQWKEFGDLSVELAHYDVLSPKNTHPNY 127
 114 -----ERGEE--SAVMLEPRHIAIKAILGLSSIGTPPEGITAEVLVYTSFDELQRRASEA 166
 128 ISIINDEGNEIFNTSLFEP-----PPGYENVSQDIVPPFSAFSPQGM-P 170
 167 RGIIVVNOPYINYSRT-----VOYRTQGAVEAAKVG-----ALASLIRSV 207
 171 EGDLYV-----VYNVARTEDFFKLERDMKINCSGKIVIAHYGVFRGNKVKNAQLAGAGV 225
 208 ASFS-----IYSPHTGIGQYOD----- 224
 226 ILYSDPADYFAP--GVKSYPDGMNLPGGGVORGNILINLNGADPLTPGYANEVAYRRGI 283
 225 -----GVKPIPTACTIVEDAEMMSMAASHGKIVIQKGAATYD----- 265
 284 AEAVGLPEIPVHPPIGYDDAQKLE-----KKGGSAPPDSSWRGSLKVPYNVGP 331
 266 -----TDSFNTVAEITGSKYEQVVLVSGHLSMDVGGAMDG 304
 332 GFTGNFSTQVKVMHISTNEVTRINVIQTLRGAVEBDRYVILGHRDSWVF--GGIDPQ 389
 305 GGAFISWE--ALSLIKDLGLRPKRTLRVLWMTAEBOGVGAFOYQOLHKVNISVSLVM 361
 390 SGAIVVHEIYVSFGTLKKEGWRPRRTILFASWDAEEFGILGSTMAE-----ENSRLLQ 443
 362 ESDAGTFLPTGLQFTGSEKARAIMIEWM-SLLQPLNTIYOVLSHGEGTDINFIQAGVGA 420
 444 ERGV-AVINADSSIEGNVTLRVDCPTLMTSLVH--NLTKELKSPD-----EGFEKG 491

QY 421 SLDDLK 428
DB 492 SLYESWTK 499

RESULT 9

US-09-044-668-2
Sequence 2, Application US/09044668
Patent No. 6150508

GENERAL INFORMATION:

APPLICANT: Murphy, Gerald P.
APPLICANT: Boynton, Alton L.
APPLICANT: Holmes, Eric H.
APPLICANT: Tino, William Thomas
TITLE OF INVENTION: MONOCLONAL ANTIBODIES SPECIFIC
TITLE OF INVENTION: FOR THE EXTRACELLULAR DOMAIN OF PROSTATE-SPECIFIC
TITLE OF INVENTION: MEMBRANE ANTIGEN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/044,668
FILING DATE: 18-MAR-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Baldwin, Geraldine F
REGISTRATION NUMBER: 31,232
REFERENCE/DOCKET NUMBER: 8511-0013-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 750 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6150508e
US-09-044-668-2

Query Match 6.8%; Score 165; DB 3; Length 750;
Best Local Similarity 19.2%; Pred. No. 2.2e-07;
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIFAFPGVHLTLSCGKAIKNGISKRTFEIKERIASCGVAKAIINLAVYKAQR 62
DB 32 FFLGLFPGWFIKSNATNTITPKNNKAFDELKAE-----NIKKFLVNF----- 77
QY 63 SYERLALLVDTVGRLSGSK--NLEKAIQIMYQNTLOQDGLKXVHLEPV-----RIPHW 113
DB 78 -----TQIPHLAGTRQNFOLAKQIOQWKEFGIDSVELAHYDVLISYPNKTIPNY 127
QY 114 -----ERGER--SAVMLEPRIRHKIALILGLSSISGTPREGITAEVLVYTSFDELQRASEA 166
DB 128 ISINEDNEIFNTSLFEP-----PPGYENVSDIVPPFSAPSQGM-P 170
QY 167 RGIIVVNOPYINVSRT-----VOYRTOGAVEAKVG-----ALASLIRSV 207

DB 171 EGDLYV-----YVYARTEDEFFKLERDMKINSGKIVARYGVFRGNKVNQAOLAGAKGV 225
QY 208 ASFS-----IYSPHTGIOEYOD----- 224
DB 226 ILVSDPADYAP--GVKSYDDGMNLPGGGVORGNILNLGAGDPLPPGYANAYARRGI 283
QY 225 -----GVKPIPTACITVEDAEKMSRMASHGIRIVITOLMGAKTYPD----- 265
DB 284 ABAVGLPSIVVHPIGYDAQKLE-----KMGSAAPDSSWRGSLKVPYVNGP 331
QY 266 -----TDSFNTVAITTSKPIPEQVVLVSGHLDSDVQOGAMDG 304
DB 332 GFTGNFSTQKXKHISTNEVTRIVYVIGTRGAVEBDRVYILGHRDSWVF--GCIDPQ 389
QY 305 GGAFLSWE--ALSLIKDGLRPRKTRILVLTAEBOGGAGFOYQOLHKVINSYSLVM 361
DB 390 SGAIVVHEIVRSQTLKESGMRPRRTILFASWPAEEGGLGSTEWE-----ENSRILQ 443
QY 362 ESDAGFLPTGLOFTGSEKARAIMIEW--SLQPLNITQVLSHGEGTDINFWIQAGVPGA 420
DB 444 ERGV-AVINADSSIEGVTIRVDCTPLMYSLVH--NUTKELKSPD-----EGREGK 491
QY 421 SLDDLK 428
DB 492 SLYESWTK 499

RESULT 10

US-08-705-477E-2
Sequence 2, Application US/08705477E
Patent No. 6569432

GENERAL INFORMATION:

APPLICANT: Israeli, Ron S
APPLICANT: Heston, Warren D.W.
APPLICANT: Fair, William R.
APPLICANT: Overfelli, Quathnek
APPLICANT: Pinto, John
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF
FILE REFERENCE: 1769/41426-G
CURRENT APPLICATION NUMBER: US/08/705,477E
CURRENT FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 750
TYPE: PRT
ORGANISM: Homo sapiens
US-08-705-477E-2

Query Match 6.8%; Score 165; DB 4; Length 750;
Best Local Similarity 19.2%; Pred. No. 2.2e-07;
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIFAFPGVHLTLSCGKAIKNGISKRTFEIKERIASCGVAKAIINLAVYKAQR 62
DB 32 FFLGLFPGWFIKSNATNTITPKNNKAFDELKAE-----NIKKFLVNF----- 77
QY 63 SYERLALLVDTVGRLSGSK--NLEKAIQIMYQNTLOQDGLKXVHLEPV-----RIPHW 113
DB 78 -----TQIPHLAGTRQNFOLAKQIOQWKEFGIDSVELAHYDVLISYPNKTIPNY 127
QY 114 -----ERGER--SAVMLEPRIRHKIALILGLSSISGTPREGITAEVLVYTSFDELQRASEA 166
DB 128 ISINEDNEIFNTSLFEP-----PPGYENVSDIVPPFSAPSQGM-P 170
QY 167 RGIIVVNOPYINVSRT-----VOYRTOGAVEAKVG-----ALASLIRSV 207
DB 171 EGDLYV-----YVYARTEDEFFKLERDMKINSGKIVARYGVFRGNKVNQAOLAGAKGV 225
QY 208 ASFS-----IYSPHTGIOEYOD----- 224
DB 226 ILVSDPADYAP--GVKSYDDGMNLPGGGVORGNILNLGAGDPLPPGYANAYARRGI 283

Db 195 -GVASYPDGMNLPGGGVORGNILNLNGADPLTPGPANEYAVRGLAEAVGLPSIVHP 253
Qy 234 ITVEDAEEMSMASHGIKIVYOLKMGATYPD----- 265
Db 254 IGYVDAQKLE-----KMGSAAPPDSSWRGLKVPYVNGPFTGNFSTOKVKM 301
Qy 266 -----TDSFNTVAEITGSKYPEQVYLVSGHLDSDVQGMADGGAFISWE---AL 314
Db 302 HIHSTNEVTRIVNYIGLRGAVEPDRYVILGHRDSWVF--GGIDPQGAHVHEIVRSF 359
Qy 315 SLIDGLRPRKTRILVMTAEBOGVGAFQYVQLHKVNISNYSLVWESDAGTFLPTGLQ 374
Db 360 GTLKKEGMRPRRTILFASMDAEFGLGSTEMAE-----ENSRLLQERGVAAYINDSS 412
Qy 375 FTGSEKARAMEVW-SILOPLNTQVLSHGEQDINFQVAGVPASLLDDLYK 428
Db 413 IEGVYTRVDCPTPLMYSLVH--NLTKEKSPD-----EGREGSLYESWTK 456

RESULT 13

PCT-US95-11720-18
; Sequence 18, Application PC/TUS9511720
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: EXPRESSION AND EXPORT TECHNOLOGY OF
; TITLE OF INVENTION: PROTEINS AS IMMUNOPROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; ADDRESSEE: THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11720
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: FIP-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7000
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 707 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..707
; OTHER INFORMATION: /note="EXTRACELLULAR DOMAIN OF
; PCT-US95-11720-18"

Query Match 6.6%; Score 160; DB 5; Length 707;

Best Local Similarity 19.6%; Pred No. 6.4e-07;
Matches 93; Conservative 60; Mismatches 144; Indels 178; Gaps 21;

Qy 76 PRLSGSK--MLEKAIQIWMYONLQODGLEKVALPEV-----RIPHW-----ERGER--S 119
Db 38 PHLAGTEQWFLAKQIQSQWKEFGLDVSLAHYVLLSYPNKTHNYISIIINEDGNELFN 97

Qy 120 AVMLEPRHIAKIALGLSSIGTPEEGITAEVLVWTSFDELQRRASARGKIVVNOFYIN 179
Db 98 TSLFEP-----PPPGYENVSIVPEPSASPSGM--PEGDLV-----YVN 135
Qy 180 YSRT-----VOYRTQAVEAAKVG-----ALASLIRSVASFS---IYSPH 216
Db 136 YARTEDEPKLERMDKINCSKIVARYGKVRGNKVNQAOLAGAKVILYSDPADYFAP- 194
Qy 217 TGIQOEYD-----GVPKIPTAC 233
Db 195 -GVASYPDGMNLPGGGVORGNILNLNGADPLTPGPANEYAVRGLAEAVGLPSIVHP 253
Qy 234 ITVEDAEEMSMASHGIKIVYOLKMGATYPD----- 265
Db 254 IGYVDAQKLE-----KMGSAAPPDSSWRGLKVPYVNGPFTGNFSTOKVKM 301
Qy 266 -----TDSFNTVAEITGSKYPEQVYLVSGHLDSDVQGMADGGAFISWE---AL 314
Db 302 HIHSTNEVTRIVNYIGLRGAVEPDRYVILGHRDSWVF--GGIDPQGAHVHEIVRSF 359
Qy 315 SLIDGLRPRKTRILVMTAEBOGVGAFQYVQLHKVNISNYSLVWESDAGTFLPTGLQ 374
Db 360 GTLKKEGMRPRRTILFASMDAEFGLGSTEMAE-----ENSRLLQERGVAAYINDSS 412
Qy 375 FTGSEKARAMEVW-SILOPLNTQVLSHGEQDINFQVAGVPASLLDDLYK 428
Db 413 IEGVYTRVDCPTPLMYSLVH--NLTKEKSPD-----EGREGSLYESWTK 456

RESULT 14

US-08-547-197-1
; Sequence 1, Application US/08547197
; Patent No. 5691157
; GENERAL INFORMATION:
; APPLICANT: Gong, Joseph K.
; APPLICANT: Glomski, Chester A.
; TITLE OF INVENTION: A METHOD FOR DETECTING A MAMMAL'S PRIOR
; TITLE OF INVENTION: EXPOSURE TO RADIATION OR RADIOMETRIC AGENTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/547,197
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rogalsky, Peter
; REGISTRATION NUMBER: 38,601
; REFERENCE/DOCKET NUMBER: 19226/580 (R-5228)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1634
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-547-197-1

Query Match 5.9%; Score 143.5; DB 1; Length 622;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:47:05 ; Search time 39.3941 Seconds
(without alignments)
2237.645 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417
Sequence: 1 MKFLIFAFGCVHLISCSG.....AWAVSYVADMEMLPRS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
 - 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
 - 6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
 - 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
 - 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
 - 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
 - 11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/ptodata/2/pubppaa/US09C_NEW_PUB.pep:*
 - 13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
 - 14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
 - 15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
 - 17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
 - 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	100.0	472	US-09-745-763-36	Sequence 36, Appl
2	2417	100.0	473	US-09-984-271-120	Sequence 120, Appl
3	188	7.8	740	US-10-325-430-15	Sequence 15, Appl
4	165	6.8	33	US-09-984-271-224	Sequence 224, Appl
5	165	6.8	34	US-09-984-271-187	Sequence 187, Appl
6	165	6.8	750	US-09-822-827-944	Sequence 944, Appl
7	165	6.8	750	US-09-895-793-944	Sequence 944, Appl
8	165	6.8	750	US-10-341-434-71	Sequence 71, Appl
9	165	6.8	750	US-10-117-937-4	Sequence 4, Appl
10	165	6.8	750	US-10-094-699-1	Sequence 1, Appl
11	165	6.8	750	US-10-205-823-132	Sequence 132, Appl
12	163	6.7	750	US-09-978-295A-618	Sequence 618, Appl
13	163	6.7	750	US-09-978-697-618	Sequence 618, Appl
14	163	6.7	750	US-09-978-192A-618	Sequence 618, Appl
15	163	6.7	750	US-09-999-832A-618	Sequence 618, Appl

16	163	6.7	750	US-09-978-189-618	Sequence 618, Appl
17	163	6.7	750	US-09-978-608A-618	Sequence 618, Appl
18	163	6.7	750	US-09-978-585A-618	Sequence 618, Appl
19	163	6.7	750	US-09-978-191A-618	Sequence 618, Appl
20	163	6.7	750	US-09-978-403A-618	Sequence 618, Appl
21	163	6.7	750	US-09-978-564A-618	Sequence 618, Appl
22	163	6.7	750	US-09-999-833A-618	Sequence 618, Appl
23	163	6.7	750	US-09-981-915A-618	Sequence 618, Appl
24	163	6.7	750	US-09-978-824-618	Sequence 618, Appl
25	163	6.7	750	US-09-918-585A-618	Sequence 618, Appl
26	163	6.7	750	US-09-978-423A-618	Sequence 618, Appl
27	163	6.7	750	US-09-978-193A-618	Sequence 618, Appl
28	163	6.7	750	US-09-999-830A-618	Sequence 618, Appl
29	163	6.7	750	US-09-978-757A-618	Sequence 618, Appl
30	163	6.7	750	US-09-978-187B-618	Sequence 618, Appl
31	163	6.7	750	US-09-978-643A-618	Sequence 618, Appl
32	163	6.7	750	US-09-978-375A-618	Sequence 618, Appl
33	163	6.7	750	US-09-978-188A-618	Sequence 618, Appl
34	163	6.7	750	US-09-978-298A-618	Sequence 618, Appl
35	163	6.7	750	US-10-143-031A-618	Sequence 618, Appl
36	163	6.7	750	US-10-002-967A-618	Sequence 618, Appl
37	163	6.7	750	US-10-017-083A-618	Sequence 618, Appl
38	163	6.7	750	US-10-143-030A-618	Sequence 618, Appl
39	163	6.7	750	US-10-199-672-104	Sequence 104, Appl
40	163	6.7	750	US-10-187-749-104	Sequence 104, Appl
41	163	6.7	750	US-10-194-457-104	Sequence 104, Appl
42	163	6.7	750	US-10-145-128A-618	Sequence 618, Appl
43	163	6.7	750	US-10-184-642-104	Sequence 104, Appl
44	163	6.7	750	US-10-196-747-104	Sequence 104, Appl
45	163	6.7	750	US-10-173-689-104	Sequence 104, Appl

ALIGNMENTS

RESULT 1
US-09-745-763-36
Sequence 36, Application US/09745763
Patent No. US20020065394A1

GENERAL INFORMATION:
APPLICANT: McCoy, Kenneth
McCoy, John M.
Lavalie, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Merberg, David
Treacy, Maurice
Spaulding, Vikki

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM

NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

```

; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 472 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: <unknown>
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-745-763-36

Query Match      100.0%; Score 2417; DB 9; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.5e-231;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIPAFGGVHLLSLCSGKAIQIMYQNLQODGLEKVLHPVRI PMMERGESB 60
DB 1 MKFLIPAFGGVHLLSLCSGKAIQIMYQNLQODGLEKVLHPVRI PMMERGESB 60
QY 61 NRSYERIALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRI PMMERGESB 120
DB 61 NRSYERIALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRI PMMERGESB 120
QY 121 VMLERPIHKIALILGSSSIGTPPEGITAEVLVTSFDELQRRASEARGKIVVYNQPIYNY 180
DB 121 VMLERPIHKIALILGSSSIGTPPEGITAEVLVTSFDELQRRASEARGKIVVYNQPIYNY 180
QY 181 SRTVQYRTQGAVEAAKVGALASLIRSVASFSTYSPHTGIOEYQDGVPKIPACTIVEDAE 240
DB 181 SRTVQYRTQGAVEAAKVGALASLIRSVASFSTYSPHTGIOEYQDGVPKIPACTIVEDAE 240
QY 241 MMSRMAHSGIKIVIQLMGAKTYPDTDSFNTVAETISGKYPEQVVLVSGHDSMDVGCGA 300
DB 241 MMSRMAHSGIKIVIQLMGAKTYPDTDSFNTVAETISGKYPEQVVLVSGHDSMDVGCGA 300
QY 301 MDDGGAFISWEALSLIKDGLRPKRTLRLVMTAEBOGGVGAFQYQLHKVNISNYSLV 360
DB 301 MDDGGAFISWEALSLIKDGLRPKRTLRLVMTAEBOGGVGAFQYQLHKVNISNYSLV 360
QY 361 MESDAGTFLPTGLQFTGSEKRAIMEEVMSLQPLNTITQVLSHGEGTDINFWIAGVPGA 420
DB 361 MESDAGTFLPTGLQFTGSEKRAIMEEVMSLQPLNTITQVLSHGEGTDINFWIAGVPGA 420
QY 421 SLDDLVKKYFFPHSHGDTMTVMDPKOMNVAANAVALVSVVADMEMLPRS 472
DB 421 SLDDLVKKYFFPHSHGDTMTVMDPKOMNVAANAVALVSVVADMEMLPRS 472

RESULT 2
US-09-984-271-120
; Sequence 120, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 120
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (473)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-120

Query Match      100.0%; Score 2417; DB 11; Length 473;
Best Local Similarity 100.0%; Pred. No. 1.5e-231;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIPAFGGVHLLSLCSGKAIQIMYQNLQODGLEKVLHPVRI PMMERGESB 60
DB 1 MKFLIPAFGGVHLLSLCSGKAIQIMYQNLQODGLEKVLHPVRI PMMERGESB 60
QY 61 NRSYERIALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRI PMMERGESB 120
DB 61 NRSYERIALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRI PMMERGESB 120
QY 121 VMLERPIHKIALILGSSSIGTPPEGITAEVLVTSFDELQRRASEARGKIVVYNQPIYNY 180
DB 121 VMLERPIHKIALILGSSSIGTPPEGITAEVLVTSFDELQRRASEARGKIVVYNQPIYNY 180
QY 181 SRTVQYRTQGAVEAAKVGALASLIRSVASFSTYSPHTGIOEYQDGVPKIPACTIVEDAE 240
DB 181 SRTVQYRTQGAVEAAKVGALASLIRSVASFSTYSPHTGIOEYQDGVPKIPACTIVEDAE 240
QY 241 MMSRMAHSGIKIVIQLMGAKTYPDTDSFNTVAETISGKYPEQVVLVSGHDSMDVGCGA 300
DB 241 MMSRMAHSGIKIVIQLMGAKTYPDTDSFNTVAETISGKYPEQVVLVSGHDSMDVGCGA 300
QY 301 MDDGGAFISWEALSLIKDGLRPKRTLRLVMTAEBOGGVGAFQYQLHKVNISNYSLV 360
DB 301 MDDGGAFISWEALSLIKDGLRPKRTLRLVMTAEBOGGVGAFQYQLHKVNISNYSLV 360
QY 361 MESDAGTFLPTGLQFTGSEKRAIMEEVMSLQPLNTITQVLSHGEGTDINFWIAGVPGA 420
DB 361 MESDAGTFLPTGLQFTGSEKRAIMEEVMSLQPLNTITQVLSHGEGTDINFWIAGVPGA 420
QY 421 SLDDLVKKYFFPHSHGDTMTVMDPKOMNVAANAVALVSVVADMEMLPRS 472
DB 421 SLDDLVKKYFFPHSHGDTMTVMDPKOMNVAANAVALVSVVADMEMLPRS 472

RESULT 3
US-10-325-430-15
; Sequence 15, Application US/10325430
; Publication No. US2003015325A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; FILE REFERENCE: MP101-294PIRNM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-325-430-15

Query Match      7.8%; Score 188; DB 12; Length 740;
Best Local Similarity 20.0%; Pred. No. 2.5e-09;
Matches 112; Conservative 85; Mismatches 186; Indels 176; Gaps 27;

QY 58 KAQN-----RSYERIALVDVTGPRLSGSKN---LEKAIQIMYQNLQODGLEKVLHP-V 108
DB 58 KAQN-----RSYERIALVDVTGPRLSGSKN---LEKAIQIMYQNLQODGLEKVLHP-V 108
```

```

Db      56 KAENIKSFLRSFTKL-----PHLAGTRONFLAKKIQTQWKKFGILDSAKLWHYDVL 107
Qy      109 RIPHMERGESASVWLSPRIKIALILGSSIGTPEGTAEVLVTSFDELQRRASBARG 168
Db      108 SYRETEANANTISIDE---HETELFKT-STLEPPDDGENTNIVPYNNAFSAQGM-PEG 162
Qy      169 KIIVVNOPIYINYSRTVOY-----RTQGAVEAAKVAL-----ASIRSVASFST 212
Db      163 DLV-----YVNVARTDEPFKLEREMGINCTGKIYIARYGKIFRGNKVKNNMLAGAIIGIL 217
Qy      213 YS-----PHTGIO-----EYQD 224
Db      218 YSPADYFAPEVOPYPRKGMILPGTAQRGNVNLNGADPLTPGYPAKEYTFLDVEBEGV 277
Qy      225 GVPIKIPACTIVEDAEKMSRMAHSGIKIVIQKMGAKT-----YPTDTSF----- 269
Db      278 GIPRIPIPHIGYNDAILKRLG-GIAPPDKSMGALANVSISIPGTGSDSPFKVMHV 336
Qy      270 -----NTVAETGSKYPEQVVLVSGHLSMDVQOGAMDGGAFISWE---ALSL 316
Db      337 YNINKIRIVNVGTIGTSVEPRRYVLGHRDSWVF--GAIPTSGAVLQELIARSPGK 394
Qy      317 IKDLGLPKRTLRLVLTAEBOGVGAFOYQLHKVINSYSLVMSDAGTFELFTGLQFT 376
Db      395 LMSKMPRRRTIIPASWDAEFGILGSTEWAEE---ENVKILQERSI-AYINSDSIE 447
Qy      377 GSEKARA-----IMEEWSLLOPL-----NITQVLSHGE 405
Db      448 GNYTLRVDCPTPLLYQLVYKLTKEIPSPDDGESKSLYESWLEKPPSPENKLPRLINKLGS 507
Qy      406 GTDINFMIQ-AGVGA-----SLDDLKYKFFHHSHGDTMTVM---DP---KOMNVA 451
Db      508 GSDEAFQRLGINSGARATYKNNKKTDXSYSPYHTIYEFELVEKFDYPTFKQLSLVA 567
Qy      452 AAWAVVSYVADMEMLP 470
Db      568 QLRGALV-YELVD-SKIIP 584

```

```

RESULT 4
US-09-984-271-224
; Sequence 224, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984, 271
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 224
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-271-224

```

```

Query Match      6.8%; Score 165; DB 11; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      440 MTWMDPKOMNVAANAVALVSYVADMEMLPRS 472
|||||

```

```

Db      1 MTWMDPKOMNVAANAVALVSYVADMEMLPRS 33
RESULT 5
US-09-984-271-187
; Sequence 187, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984, 271
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 187
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-187

```

```

Query Match      6.8%; Score 165; DB 11; Length 34;
Best Local Similarity 100.0%; Pred. No. 3.3e-09;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      440 MTWMDPKOMNVAANAVALVSYVADMEMLPRS 472
|||||
Db      1 MTWMDPKOMNVAANAVALVSYVADMEMLPRS 33
|||||

```

```

RESULT 6
US-09-822-827-944
; Sequence 944, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 944
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-944

```

```

Query Match      6.8%; Score 165; DB 9; Length 750;
Best Local Similarity 19.2%; Pred. No. 5e-07;
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;
Qy      3 FLIAFGVYHLISLGSKAICKXIGSKPTPEEIKESIASGVAKAIVMLAVGKAQNR 62
|||||
Db      32 FFLGLFPGWFISSNENATITTPAHNNKAFIDELKAE-----NKKFLVNF----- 77
|||||
Qy      63 SYERIALLVTVGPRLSGSK---NLEKAIQIMYONLQODGLEKVTLPV-----RIPHW 113
|||||

```

Db 78 -----TQIPLHAGTEQNFOLAKQIOSQWKEFGDVELAHYDVLLSYPNKTHPNY 127
Qy 114 -----ERGEE--SAVMELEPRHKAIIILGSSICTPPGIGTAELVYVTSFDELGRASEA 166
Db 128 ISIINEDNEIENFNTSLFEP-----PPPYENVSIDIVPPSASPDQM-P 170
Qy 167 RGIIVVYQPYIYNSRT-----VOYRTQGAWEAAKV-----ALASLIRSV 207
Db 171 EGDLY-----YVYAKTDEPFKLEMDKINKSGKIIVARYGVFRGNKVKQAOLAGAKGV 225
Qy 208 ASFS-----TSPHTGIGYOD----- 224
Db 226 ILVSPDADYFAP--GVKSYPDGMNLPGGVQGRNIIILNAGADPLTPGPANAYARRGI 283
Qy 225 -----GVKIPACTIVDAEMSRMASHGIKIIVQLMGAKTYPD----- 265
Db 284 AEAVGLPSIPVHPIDIGYDAQKLE-----KMGGSAPPDSSWRGSLKVPYVGP 331
Qy 266 -----TDSFNTVAEITGSKYPEQVVLVSGHLDSDVGGAMDG 304
Db 332 GFTGNFSTQKVKMHISTNEVTRINVTIGTLRGAVEBDRVYILGHRDSWF--GGIDPQ 389
Qy 305 GGAFISWE--ALSILKDLGRPKRTLRLVMTAEEOGVGAFOYQLHKVNISVSLW 361
Db 390 SGAIVVHEIYVSFGTLKKEGWRPRTILFASWDAEEFGILGSTEWAE-----ENSRLQ 443
Qy 362 ESDAGTFLPTGLQTFGSEKARAIMIEW-SILOPLNITOVLSHGEGTDINFIQAGVGA 420
Db 444 ERGV-AYINADSIIEGNYTLRVDCPTLPMYSLVH--NLTKELKSPD-----EGFEKG 491
Qy 421 SLDDLYK 428
Db 492 SLYESWTK 499

RESULT 7

US-09-895-793-944

Sequence 944, Application US/09895793

Publication No. US20020192763A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiaqichun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiaq, Yuqiu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedrick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yahir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Basbols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.534C2
CURRENT APPLICATION NUMBER: US/09/895,793
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 944
LENGTH: 750
TYPE: PRT
ORGANISM: Homo sapiens

US-09-895-793-944

Query Match 6.8%; Score 165; DB 10; Length 750;

Best Local Similarity 19.2%; Pred. No. 5e-07;

Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

Qy 3 FLIAPFGVYHLSLSCGKAICKNGISKRTPEEIKERIASCGVAKAIIINLAVYGAQNR 62
Db 32 FFLGLFEGMVKISNENATNITPKNMKAPLDELKAE-----NKKFLYNF----- 77
Qy 63 SYERLALLVDTGVRLSGSK--NLEKAIQIMYONLOODEKYNLEPV-----RIPHW 113
Db 78 -----TQIPLHAGTEQNFOLAKQIOSQWKEFGDVELAHYDVLLSYPNKTHPNY 127
Qy 114 -----ERGEE--SAVMELEPRHKAIIILGSSICTPPGIGTAELVYVTSFDELGRASEA 166
Db 128 ISIINEDNEIENFNTSLFEP-----PPPYENVSIDIVPPSASPDQM-P 170
Qy 167 RGIIVVYQPYIYNSRT-----VOYRTQGAWEAAKV-----ALASLIRSV 207
Db 171 EGDLY-----YVYAKTDEPFKLEMDKINKSGKIIVARYGVFRGNKVKQAOLAGAKGV 225
Qy 208 ASFS-----TSPHTGIGYOD----- 224
Db 226 ILVSPDADYFAP--GVKSYPDGMNLPGGVQGRNIIILNAGADPLTPGPANAYARRGI 283
Qy 225 -----GVKIPACTIVDAEMSRMASHGIKIIVQLMGAKTYPD----- 265
Db 284 AEAVGLPSIPVHPIDIGYDAQKLE-----KMGGSAPPDSSWRGSLKVPYVGP 331
Qy 266 -----TDSFNTVAEITGSKYPEQVVLVSGHLDSDVGGAMDG 304
Db 332 GFTGNFSTQKVKMHISTNEVTRINVTIGTLRGAVEBDRVYILGHRDSWF--GGIDPQ 389
Qy 305 GGAFISWE--ALSILKDLGRPKRTLRLVMTAEEOGVGAFOYQLHKVNISVSLW 361
Db 390 SGAIVVHEIYVSFGTLKKEGWRPRTILFASWDAEEFGILGSTEWAE-----ENSRLQ 443
Qy 362 ESDAGTFLPTGLQTFGSEKARAIMIEW-SILOPLNITOVLSHGEGTDINFIQAGVGA 420
Db 444 ERGV-AYINADSIIEGNYTLRVDCPTLPMYSLVH--NLTKELKSPD-----EGFEKG 491
Qy 421 SLDDLYK 428
Db 492 SLYESWTK 499

RESULT 8

US-10-341-434-71

Sequence 71, Application US/10341434

Publication No. US20030215835A1

GENERAL INFORMATION:

APPLICANT: Origene Technologies
TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
FILE REFERENCE: 9U 204 205 R1
CURRENT APPLICATION NUMBER: US/10/341,434
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: US 60/348,164
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: US 60/348,119
PRIOR FILING DATE: 2002-01-15
NUMBER OF SEQ ID NOS: 238
SOFTWARE: PatentIn version 3.1
SEQ ID NO 71
LENGTH: 750
TYPE: PRT
ORGANISM: Homo sapiens
US-10-341-434-71

Query Match 6.8%; Score 165; DB 12; Length 750;
Best Local Similarity 19.2%; Pred. No. 5e-07;
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

Query Match	6.8%	Score 165;	DB 12;	length 750;
Best Local Similarity	19.2%;	Pred. No. 5e-07;		
Matches 105;	Conservative 67;	Mismatches 174;	Indels 202;	Gaps 23;

Query Match 6.8%; Score 165; DB 15; Length 750;
Best Local Similarity 19.2%; Pred. No. 5e-07;
Matches 105; Conservative 67; Mismatches 174; Gaps 23;

QY 3 FLFAFGGHHLLSCGALICNGISRTPEIREEELASGVDAAKALINAYGKQNR 62
Db 32 FFLFLGFLGFWIKSSNEATNITPKANMAFLDELQAE-----NIKFLYNF----- 77
QY 63 SYERIALVDTVPRSSGK---NLEKALQIMYCNLQDDGLKXHLBEV-----RIPHW 113

Db 78 -----TOIPLHAGTEQNFOLAKOIOSQWKEFGDLSVELAHYDVLISYPNKTHPNY 127
Qy 114 -----ERGEE--SAVMEPRRIHKAIILGLSSICTPPEGITAEVLYVTSFDELQRRASEA 166
Db 128 ISIINEDENETFNISLSEF-----PPPEYENSDIVPPFSAFSPQGM-P 170
Qy 167 RGIIVVYNQPIYINSRT-----VOYRTOGAWEAAKV-----ALASLIRSV 207
Db 171 EGDLY-----YVYAKTDEPFKLERDMKINCSCGKIIVARIGKVPKKNKNAQLAGAKGV 225
Qy 208 ASFS-----IYSPHTGIEYOD----- 224
Db 226 ILYSDPADYFAP--GVKSYPDGMNLPGGGVQGRNIIILNAGADPLTPGPANEYAYRRGI 283
Qy 225 -----GVKPIPTACTVDEDAEMSRMAHSGIKIIVILKMGAKTYPD----- 265
Db 284 AEAVGLPSIPVHPHIGYIDQKLE-----KMGGSAPPDSSWRGSLKVPYVGP 331
Qy 266 -----TDSFNTVAETGSKYPEQVVLVSGHLSMDVQGMDDG 304
Db 332 GFTGNFSTOKVMHISTNEVTRIVNIGTLRGAVEPRRYVILGHRDSWVF--GGLDPQ 389
Qy 305 GGAFISWE--ALSILKDLGLRPKRTLVLVMTAEQGVGAFOYYQLHKVNISNYSLVW 361
Db 390 SGAIVVHEIVRSFGLTKKEGWRPRTILFASWDAEEFGLGSTEWAE-----ENSRLLQ 443
Qy 362 ESDAGTFLPTGLQTFGSEKARAIMIEW--SLLOPLNITOVLSHGCTINFWIQAGVGA 420
Db 444 ERGV-AYINADSSIEGNTYLRVDCPTLPMYSLVH--NLTKELKSPD-----EGFBGK 491
Qy 421 SLDDLYK 428
Db 492 SLYESWTK 499

RESULT 11
US-10-205-823-132
; Sequence 132, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gothatcheva, Bella
; APPLICANT: Hoerscn, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatc, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 132
LENGTH: 750
TYPE: PRT
ORGANISM: Homo sapiens

US-10-205-823-132

Query Match 6.8%; Score 165; DB 15; Length 750;
Best Local Similarity 19.2%; Pred. No. 5e-07;
Matches 105; Conservative 67; Mismatches 114; Indels 202; Gaps 23;

Qy 3 FLIAPFGVYHLSLSCGKAICKNGISKRTFEIEKEIEIACGDAKAIINLAAYGKAQR 62
Db 32 PFLGLPLFGWFIKSNATNITTPHNNKAPLDELKAE-----NIKFLYNF----- 77
Qy 63 SYERLLALVDTVGRISGSK--NLEKAQIMQNTLOODGLEKYHLEPV-----RIPHW 113
Db 78 -----TOIPLHAGTEQNFOLAKOIOSQWKEFGDLSVELAHYDVLISYPNKTHPNY 127
Qy 114 -----ERGEE--SAVMEPRRIHKAIILGLSSICTPPEGITAEVLYVTSFDELQRRASEA 166
Db 128 ISIINEDENETFNISLSEF-----PPPEYENSDIVPPFSAFSPQGM-P 170
Qy 167 RGIIVVYNQPIYINSRT-----VOYRTOGAWEAAKV-----ALASLIRSV 207
Db 171 EGDLY-----YVYAKTDEPFKLERDMKINCSCGKIIVARIGKVPKKNKNAQLAGAKGV 225
Qy 208 ASFS-----IYSPHTGIEYOD----- 224
Db 226 ILYSDPADYFAP--GVKSYPDGMNLPGGGVQGRNIIILNAGADPLTPGPANEYAYRRGI 283
Qy 225 -----GVKPIPTACTVDEDAEMSRMAHSGIKIIVILKMGAKTYPD----- 265
Db 284 AEAVGLPSIPVHPHIGYIDQKLE-----KMGGSAPPDSSWRGSLKVPYVGP 331
Qy 266 -----TDSFNTVAETGSKYPEQVVLVSGHLSMDVQGMDDG 304
Db 332 GFTGNFSTOKVMHISTNEVTRIVNIGTLRGAVEPRRYVILGHRDSWVF--GGLDPQ 389
Qy 305 GGAFISWE--ALSILKDLGLRPKRTLVLVMTAEQGVGAFOYYQLHKVNISNYSLVW 361
Db 390 SGAIVVHEIVRSFGLTKKEGWRPRTILFASWDAEEFGLGSTEWAE-----ENSRLLQ 443
Qy 362 ESDAGTFLPTGLQTFGSEKARAIMIEW--SLLOPLNITOVLSHGCTINFWIQAGVGA 420
Db 444 ERGV-AYINADSSIEGNTYLRVDCPTLPMYSLVH--NLTKELKSPD-----EGFBGK 491
Qy 421 SLDDLYK 428
Db 492 SLYESWTK 499

RESULT 12
US-09-78-295A-618
; Sequence 618, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gettisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijaviri, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James J.
; APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann
 APPLICANT: Shelton, David L.
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: P2630P1C11
 CURRENT APPLICATION NUMBER: US/09/978,295A
 CURRENT FILING DATE: 2001-10-15
 PRIOR APPLICATION NUMBER: 09/918585
 PRIOR FILING DATE: 2001-07-30
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/064249
 PRIOR FILING DATE: 1997-11-03
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066364
 PRIOR FILING DATE: 1997-11-21
 PRIOR APPLICATION NUMBER: 60/077450
 PRIOR FILING DATE: 1998-03-10
 PRIOR APPLICATION NUMBER: 60/077632
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077641
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077649
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077791
 PRIOR FILING DATE: 1998-03-12
 PRIOR APPLICATION NUMBER: 60/078004
 PRIOR FILING DATE: 1998-03-13
 PRIOR APPLICATION NUMBER: 60/078886
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078936
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078939
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/079294
 PRIOR FILING DATE: 1998-03-25
 PRIOR APPLICATION NUMBER: 60/079656
 PRIOR FILING DATE: 1998-03-26
 PRIOR APPLICATION NUMBER: 60/079664
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079689
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079663
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079728
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079786
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079920
 PRIOR FILING DATE: 1998-03-30
 PRIOR APPLICATION NUMBER: 60/079923
 PRIOR FILING DATE: 1998-03-30
 PRIOR APPLICATION NUMBER: 60/080105
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080107
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080165
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080194
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080327
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080328
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080333
 PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/080334
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/081070
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081049
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081071
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081195
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081955
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081819
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081952
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081838
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082568
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082569
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082704
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082804
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082700
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082797
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082796
 PRIOR FILING DATE: 1998-04-23
 PRIOR APPLICATION NUMBER: 60/083336
 PRIOR FILING DATE: 1998-04-27
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083392
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083495
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083496
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083499
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083554
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083558
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083559
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083500
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083742
 PRIOR FILING DATE: 1998-04-30
 PRIOR APPLICATION NUMBER: 60/084366
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084441
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640

PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.7% Score 163; DB 10; Length 750;

Best Local Similarity 19.2%; Pred. No. 7, 9e-07; Indels 202; Gaps 23;
Matches 105; Conservative 67; Mismatches 174;

QY 3 FLIFAFGVLHSLCSGKAIKNGISKRTPEEIKETIASCGDVAKAIINLAVGKAQR 62
DB 32 FLIFGFLGFWFIKSNENATNITPKNNMFAFDELAF-----NIKKFLHNF----- 77
QY 63 SYERLALLVDYVGPLSGSK--NIEKAIQIMYQNLQODGLEKVLLEV-----RIHW 113
DB 78 -----TQIPHLAGEQNFQAKQIQSQWKEFGDSEVLAHYDVLISYNNKTHNY 127
QY 114 -----ERGEE--SAVMEPRRIHKIAILGLSSIGPPEGITAEVLVYMSFDELQRASEA 166
DB 128 ISIINEDGNEIFNISLFEP-----PPGEYENVSQIVPPFSAFSPQGM-P 170
QY 167 RGIIVVYQPYINVSRT-----VOYRTGAVEAAKVG-----ALASLIRSV 207
DB 171 BGDLYV-----YVNVARTEDPFLERDMKINSGKIVIRYGVFRGNKVKNAQLAGAGV 225
QY 208 ASFS-----TSPHTSIGQYOD----- 224
DB 226 ILVSPADYFAP--GVKSYPDGMNLPGGGVQGRNIIINLNGAGDPYTPGPAVEYVRRGI 283
QY 225 -----GVKIPRACTIVEDEAMMSMAHSHGIRKIVILKMGAKTYPD----- 265
DB 284 AEAVGLPSIIVHPPIGYDAQKLB-----KKGGAAPPSVSRGSLKVPYVNP 331
QY 266 -----TDSFNTVAETGSKYPEQVVLVSGHLSWPDGQAMDG 304
DB 332 GFTGNFQKVMHISTHNEVTRINVGITLGAVERPRVYVILGHRDSWF--CGIDPQ 389
QY 305 GGAFTSW--ALSLIKDLGLRPKKTLLVMTAEQGVGAFQYYQLHKVNISVSLVM 361
DB 390 SGAAVVHIVRSFGLTKKEGMPRTIIFASWDAEEFGLGSTEWAE-----ENSRLIQ 443
QY 362 ESDAGTFLPTGQFGSGEKARAIMENV--SLQPINIYOVLSHGSGTINIMIQGVGA 420
DB 444 ERGVAYINADSSIEGNYTLRDCTPLMYSLVH--NLTKELKSPD-----EGFEKG 491
QY 421 SLDDLK 428
|| : |

DB 492 SLYESWTK 499
RESULT 13
US-09-978-697-618
Sequence 618, Application US/09978697
Patent No. US20020169284A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Deenoyers, Luc
APPLICANT: Baton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C27
CURRENT FILING DATE: 2001-10-16
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656

;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29

;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.7%; Score 163; DB 10; Length 750;
Best Local Similarity 19.2%; Pred. No. 7.9e-07;
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

Qy 3 FLTAFPGVHLSTLSCGKAIKNGISKRTFEIEKIEIASCQDVAKAIIMLAVYKAQNR 62
Db 32 FPLGLFPGWPKISSNATNITPRHNNKAPFLDELKAE-----NKKFLHNF----- 77
Qy 63 SYERLALLVDTGVRSLGSK---NLEKAIQIMTONLQODGLEKVNHEPV-----RIPHW 113
Db 78 -----TQIPHLAGTEONFQALAKOIOSOMWFEGLDVEVLAAHYVLLSYPKKTHPNY 127
Qy 114 -----ERGEE--SAVMLEPRTHKIALIIGLSSICTPREGITAEVLVVTSPDELOGRASEA 166

Db 128 ISIINEDGNEIFNTSLFEP-----PPRCYENVSDIVPFSAFSPQM-P 170
Qy 167 RGIIVVNPPIYINERT-----VOYRIGAVNAKVG-----ALASLISV 207
Db 171 EGDIV-----YVNVARTDEFFLERDMKINCSGKIYIARYGVFRGNVKNQALGAGV 225
Qy 208 ASFS-----IYSPHTIOEXOD----- 224
Db 226 ILVSPPIYFAP--GVKSTPDCMNIPGGGVORGNILINAGADPLITPGYRANVAYRGI 283
Qy 225 ---GVKPIPTACTVEDAEMMSRMAHGIKIVIQKXGAKTYPD----- 265
Db 284 AEAVALPSIPVHPVIGYDAQKLE-----KKGSGAPDSSWRGSLKVPYVNGP 331
Qy 266 -----TDSFNVAETGSKYPEQVYLVSHLHLSMVDGAMDDG 304
Db 332 GFTGNFSTQKVMKIHSTNEVTRINVIQTLGAVEPRRYILLGHRDSWPF--GGIDPQ 389
Qy 305 GGAFTSWE--ALSLIKDILGPRKRTLRLVLTAEQGVGAFOYYOLHKVNISYSLVM 361
Db 390 SGAAVHVHIVSRFGLKKEGMRPRRTILFASMDAEFFGLISTEMAF-----ENSRILQ 443
Qy 362 ESDACTPLPTGLOFTGSEKARAMEVM-SLLOPLNITQVLSHGEGTDINFWIQAGVGA 420
Db 444 ERGV-AVINADSSIGNVTLRVDCTPLMYSLVH--NLTFELKSPD-----EGFEGK 491
Qy 421 SLUDDLK 428
Db 492 SLVESWTK 499

RESULT 14

US-09-978-192A-618

Sequence 618, Application US/09978192A

Patent No. US20020177553A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi

APPLICANT: Baker Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan

APPLICANT: Ferrara, Napoleon

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Kijavlin, Ivar J.

APPLICANT: Kuo, Sophia S.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Shelton, David L.

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

Acids Encoding the Same

FILE REFERENCE: P263091C9

CURRENT APPLICATION NUMBER: US/09/978,192A

CURRENT FILING DATE: 2001-10-15

PRIOR APPLICATION NUMBER: 09/918585

PRIOR FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

;; PRIOR APPLICATION NUMBER: 60/064249
;; PRIOR FILING DATE: 1997-11-03
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066364
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/077450
;; PRIOR FILING DATE: 1998-03-10
;; PRIOR APPLICATION NUMBER: 60/077632
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077641
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077649
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/078004
;; PRIOR FILING DATE: 1998-03-13
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078936
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955

APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PLC63
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/07450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30

PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500

PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-11
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.7%; Score 163; DB 10; Length 750;
Best Local Similarity 19.2%; Pred. No. 7.9e-07;
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIFAFGCVHLISCSGKAIKNGISKRTFBEIKBEIASCGDVAKAIINLAVYKQNR 62
DB 32 FFLIGFLFGWFIKSNENATITPKKMKAFIDELKAE-----NIKKFLHNF----- 77
QY 63 SYERIALLVDTVGPRLSSK--NLEKAIQIMYONTQQDGLKXVHLEPV-----RIPHW 113
DB 78 -----TQIPHLAGEONFQAKQIQSQWKEFGDSVELAHYDVLISYPNKTHPNY 127
QY 114 -----ERGEF--SAVMEPRRIHKIAILGLSSIGTPPEGCITAEVLVVTSPDELQRRASEA 166
DB 128 ISIINEDENEIFNFSLEFEP-----PPGYENVSIDIVPPFSAFSPQGM-P 170
QY 167 RGTIVVYVQPIYNSRT-----VOYRTQGAVEAKVG-----ALASLIRSV 207
DB 171 EGDLV-----YVNTARTEDPFKERDMKINCSGKIIVIRYGVFRGNVKNVNAQJLAGAKGV 225
QY 208 ASFS-----IYSPHTGIOEYOD----- 224
DB 226 ILVSDPADYFAP--GVKSYPOGMNLPGGVORGNILNLNGAGDPLTPGYPANEXAYRRGI 283
QY 225 -----GVKPIPTACTIVEDAEKMSRMAHSGIKIVIQKMGAKTYPD----- 265
DB 284 AEAVALPSIIPVPIGYDAOKLLE-----KMGSAAPDSSWRGSLKVPYVNVGP 331

QY 266 -----TDSFNTVAEITGSKYPEQVVLVSGHLSNDVQGAMDDG 304
DB 332 GFTGNFSTQKVKMHISTNEVTIRIYVIGTLRGAVEPDRYVILGHRDSWVF--GGIDPQ 389
QY 305 GGAFISWE--ALSIXDGLRPRKTLRLVLTAEBOGGVGAQYVQLHKVNISNLSVM 361
DB 390 SGAIVHEIVRSFGLTKKEGMRPRRTILFASMDAEFEGLGISTEMAE-----ENSRLQ 443
QY 362 ESDAGTFLPTGLQFTGSEKARAIMEEVM--SLQPLNITOVLSHGEGTDINFWIQAGVPGA 420
DB 444 ERGV-AYINADSSIEGNYTLRVDCPTPLMYSLVH--NLTKEKSPD-----EGFEKG 491
QY 421 SLDDLYK 428
DB 492 SLYESWTK 499

Search completed: December 22, 2003, 11:54:05
Job time: 42.3941 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:34:19 ; Search time 29.7527 Seconds

(without alignments)
933.601 Million cell updates/sec

Title: US-09-745-763-36_COPY_1_175

Perfect score: 883

Sequence: 1 MKELIFAFGFGVHLSTSCG.....FDEIQRRAEPARKIVVYQ 175

Scoring table: BL0SIM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_19Jun03.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	883	100.0	472	18	AAW33604
2	883	100.0	472	19	AAW85456
3	883	100.0	472	21	AAAB10229
4	883	100.0	472	21	AAAB10229
5	883	100.0	472	21	AAAB10229
6	883	100.0	472	21	AAAB10229
7	883	100.0	472	21	AAAB10229
8	883	100.0	472	21	AAAB10229
9	883	100.0	472	21	AAAB10229

10	883	100.0	472	24	ABR48106
11	883	100.0	472	24	ABR00152
12	883	100.0	472	24	ABR00274
13	883	100.0	472	21	AAAB7081
14	883	100.0	472	21	ABG09405
15	883	100.0	472	21	AAAB7081
16	883	100.0	472	21	AAAB7081
17	883	100.0	472	21	AAAB7081
18	883	100.0	472	21	AAAB7081
19	883	100.0	472	21	AAAB7081
20	883	100.0	472	21	AAAB7081
21	883	100.0	472	21	AAAB7081
22	883	100.0	472	21	AAAB7081
23	883	100.0	472	21	AAAB7081
24	883	100.0	472	21	AAAB7081
25	883	100.0	472	21	AAAB7081
26	883	100.0	472	21	AAAB7081
27	883	100.0	472	21	AAAB7081
28	883	100.0	472	21	AAAB7081
29	883	100.0	472	21	AAAB7081
30	883	100.0	472	21	AAAB7081
31	883	100.0	472	21	AAAB7081
32	883	100.0	472	21	AAAB7081
33	883	100.0	472	21	AAAB7081
34	883	100.0	472	21	AAAB7081
35	883	100.0	472	21	AAAB7081
36	883	100.0	472	21	AAAB7081
37	883	100.0	472	21	AAAB7081
38	883	100.0	472	21	AAAB7081
39	883	100.0	472	21	AAAB7081
40	883	100.0	472	21	AAAB7081
41	883	100.0	472	21	AAAB7081
42	883	100.0	472	21	AAAB7081
43	883	100.0	472	21	AAAB7081
44	883	100.0	472	21	AAAB7081
45	883	100.0	472	21	AAAB7081

ALIGNMENTS

RESULT 1	AAW33604	AAW33604 standard; Protein; 472 AA.
AC	AAW33604;	
XX	21-MAY-1998 (first entry)	
XX	Human secreted protein AM282 full-length sequence.	
XX	Secreted protein; AM282; cytokine; human.	
XX	Homo sapiens.	
XX	Key	Location/Qualifiers
XX	Peptide	1..24
XX	Protein	/label= sig_peptide
XX		25..472
XX		/label= Mat_protein
XX	WO9739030-A2.	
XX	23-OCT-1997.	
XX	16-APR-1997;	97WO-US06475.
XX	13-JAN-1997;	97US-0783520.
XX	18-APR-1996;	96US-0634325.
XX	(GENY) GENETICS INST INC.	
XX	Jacobs K, LaVallie ER, McCoy JM, Werberg D, Racie LA;	

PI Spaulding V;
 XX
 DR WPI: 1997-526400/48.
 DR N-PSDB; AAV02296.
 XX
 PT New isolated secretory proteins AM340, AM282 and AK583 - possibly
 PT have cytokine, cell proliferation/differentiation regulating,
 PT immunomodulating activities, etc.
 XX
 PS Claim 16; Page 45-47; 59pp; English.
 XX
 CC This human secreted protein, designated AM282, is encoded by a
 CC full-length cDNA clone (see AAV02296), deposited in ATCC 98026, that
 CC was identified from a database search using an isolated partial
 CC AM282 clone (see AAT97398). AM282 protein can be used in a claimed
 CC method for preventing, treating or ameliorating a medical
 CC condition. It may exhibit cytokine, cell proliferation (either
 CC inducing or inhibiting) or cell differentiation (either inducing or
 CC inhibiting) activity or may induce production of other cytokines in
 CC certain cell populations. It may also exhibit e.g. immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic or chemokinetic activity, haemostatic or thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC tumour inhibition activity, or other activities. No evidence of
 CC any of these activities is given in the specification.
 XX
 SQ Sequence 472 AA;
 Query Match 100.0%; Score 883; DB 18; Length 472;
 Best Local Similarity 100.0%; Pred. No. 2,2e-83;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTFEIEKEIASCGDVAKAIINLAAYGKAQ 60
 Db 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTFEIEKEIASCGDVAKAIINLAAYGKAQ 60
 Oy 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYONLQDDGLEKVLPEVRIPIHMERGESA 120
 Db 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYONLQDDGLEKVLPEVRIPIHMERGESA 120
 Oy 121 VMLERPIHKIAIILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASEARGKIVVYNQ 175
 Db 121 VMLERPIHKIAIILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASEARGKIVVYNQ 175

RESULT 2
 AAW85456
 ID AAW85456 standard; Protein; 472 AA.
 XX
 AC AAW85456;
 XX
 DT 25-FEB-1999 (first entry)
 XX
 DE Secreted protein encoded by clone bu45_2.
 XX
 KW Secreted protein; nutritional activity; immune stimulating; vaccine;
 KW suppressing activity; haematopoiesis regulating activity;
 KW tissue growth activity; activin; inhibin activity; chemotactaxis;
 KW chemokinetic activity; haemostasis; thrombolytic activity; receptor;
 KW ligand; anti-inflammatory; cadherin; tumour invasion suppressor;
 KW tumour inhibition; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9842739-A2.
 XX
 PD 01-OCT-1998.
 XX
 PF 20-MAR-1998; 98WO-US05653.
 XX
 PR 19-MAR-1998; 98US-0044466.
 PR 21-MAR-1997; 97US-0822167.

XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie LA, Spaulding V, Treacy M;
 XX
 DR WPI: 1998-609890/51.
 DR N-PSDB; AAV82779.
 XX
 PT New polynucleotides encoding secreted human proteins - derived from
 PT human foetal brain, adult brain, foetal kidney, placenta or adult
 PT pineal gland cDNA libraries.
 XX
 PS Claim 14; Page 70-72; 113pp; English.
 XX
 CC The present sequence represents a secreted protein. The polynucleotide
 CC and secreted protein are predicted to have biological activities which
 CC would make them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is given.
 CC Suggested activities include nutritional activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity (no data is
 CC given in the specification to support these activities). The
 CC polynucleotide is also stated to be useful for gene therapy.
 XX
 SQ Sequence 472 AA;
 Query Match 100.0%; Score 883; DB 19; Length 472;
 Best Local Similarity 100.0%; Pred. No. 2,2e-83;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTFEIEKEIASCGDVAKAIINLAAYGKAQ 60
 Db 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTFEIEKEIASCGDVAKAIINLAAYGKAQ 60
 Oy 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYONLQDDGLEKVLPEVRIPIHMERGESA 120
 Db 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYONLQDDGLEKVLPEVRIPIHMERGESA 120
 Oy 121 VMLERPIHKIAIILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASEARGKIVVYNQ 175
 Db 121 VMLERPIHKIAIILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASEARGKIVVYNQ 175

RESULT 3
 AAB10229
 ID AAB10229 standard; Protein; 472 AA.
 XX
 AC AAB10229;
 XX
 DT 16-NOV-2000 (first entry)
 XX
 DE Human fetal kidney protein fragment AM282_11.
 XX
 KW Secreted protein; cytostatic; immunostimulatory; antimicrobial;
 KW antiviral; immunosuppressive; antiinflammatory; vulnery; cytokine;
 KW cell proliferation; differentiation; regulator; treatment; tumor;
 KW autoimmune disease; inflammatory disorder; wound; microbial infection;
 KW viral disease; graft versus host reaction suppression.
 XX
 OS Homo sapiens.
 XX
 PN WO200037630-A1.
 XX
 PD 29-JUN-2000.
 XX
 PF 22-DEC-1999; 99WO-US31005.
 XX
 PR 23-DEC-1998; 98US-0220876.

PA (GENY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR;
 XX
 DR WPI: 2000-442661/38.
 DR N-PSDB: AAA40493.
 XX
 PT Secreted human proteins AS296-11 and AS34-11, useful for treating
 PT tumors, autoimmune diseases, inflammatory disorders, wounds, microbial
 PT infections and viral diseases -
 XX
 PS Disclosure; Page 199-200; 293pp; English.
 XX
 CC This invention describes novel secreted human proteins (I) which have
 CC cytostatic, immunostimulatory, antimicrobial, antiviral,
 CC immunosuppressive, anti-inflammatory and antiviral activity and which act
 CC as cytokine, cell proliferation or differentiation regulators. (I)
 CC is useful for treating tumors, autoimmune diseases, inflammatory
 CC disorders, wounds, microbial infections and viral diseases. (I) is also
 CC useful for suppressing graft versus host reaction. AAB10226-B10288
 CC represent the secreted proteins encoded by AAA40490-A40580 which are
 CC described in the method of the invention.
 CC
 SQ Sequence 472 AA;
 XX
 Query Match 100.0%; Score 883; DB 21; Length 472;
 Best Local Similarity 100.0%; Pred. No. 2.2e-83;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKFLIFAFPGVHLLSLCSGKAIKNGISKRTPEEIKKEELASCGDVAKAIINLAIVGKAQ 60
 DB 1 MKFLIFAFPGVHLLSLCSGKAIKNGISKRTPEEIKKEELASCGDVAKAIINLAIVGKAQ 60
 QY 61 NRSYERLALVDYVGPRLSSKNLEKAIQIMYQNLQDGLKYLEPVRIPHERGESB 120
 DB 61 NRSYERLALVDYVGPRLSSKNLEKAIQIMYQNLQDGLKYLEPVRIPHERGESB 120
 QY 121 VMLEPRHKAIALIGLSSIGTPPEGITAEVLVVTSPDELORRASEARGKIVVYNO 175
 DB 121 VMLEPRHKAIALIGLSSIGTPPEGITAEVLVVTSPDELORRASEARGKIVVYNO 175
 DE 11-MAY-2000 (first entry)
 XX
 XX Human signal peptide containing protein HSP-31 SEQ ID NO:31.
 KW Human; signal peptide-containing protein; HSP; diagnosis; cancer;
 KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;
 KW antimicrobial; neuroprotective; cardiovascular; hepatotropic;
 KW antiaesthetic; gene therapy; cell proliferation; neurological disorder;
 KW reproductive disorder; developmental disorder; arteriosclerosis;
 KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
 KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
 KW Parkinson's disease; Huntington's disease; ovulatory defect;
 KW muscular dystrophy.
 KM
 XX
 OS Homo sapiens.
 XX
 PN WO200000610-A2.
 XX
 PD 06-JAN-2000.
 XX
 PF 25-JUN-1999; 99WO-US14484.
 XX
 PF 26-JUN-1998; 98US-0090762.
 XX
 PR 31-JUL-1998; 98US-0094983.

PR 01-OCT-1998; 98US-0102686.
 PR 11-DEC-1998; 98US-0112129.
 XX
 PA (INCYTE) INCYTE PHARM INC.
 XX
 PI Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;
 PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
 PI Bandman O;
 XX
 DR WPI: 2000-160673/14.
 DR N-PSDB: AA298139.
 XX
 PT New human signal peptide-containing proteins useful in treatment,
 PT prevention and diagnosis of e.g. cancer, inflammation and
 PT cardiovascular disease -
 XX
 PS Claim 1; Page 179-180; 327pp; English.
 XX
 CC AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the
 CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have
 CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,
 CC neuroprotective, cardiovascular and antiaesthetic activities, and can
 CC be used in gene therapy. HSPs can be used to treat or prevent disorders
 CC associated with decreased activity or function of HSP. Antagonists of
 CC HSP are used to treat or prevent disorders associated with increased
 CC activity or function of HSP. Such diseases include cell proliferation
 CC (including cancer), inflammation, cardiovascular, neurological,
 CC reproductive or developmental disorders, (e.g. arteriosclerosis,
 CC cirrhosis, psoriasis, acquired immune deficiency syndrome, congestive or
 CC asthma, Crohn's disease, microbial or other infections, anaemia, or
 CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
 CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP
 CC nucleic acids can be used for the recombinant production of HSP, for
 CC detecting HSP in standard hybridisation and amplification assays (for
 CC diagnosis and monitoring), in gene therapy, as antisense,
 CC triplex-forming or ribozyme therapeutics, for detecting related sequences
 CC or genetic variations, and for chromosomal mapping. HSP are also used to
 CC raise specific antibodies (Ab) and to screen for agonists and
 CC antagonists (potential therapeutic agents). Ab are used to diagnose, or
 CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic
 CC antagonists, in competitive drug screens, and for purification of HSP
 CC from natural sources.
 CC
 SQ Sequence 472 AA;
 XX
 Query Match 100.0%; Score 883; DB 21; Length 472;
 Best Local Similarity 100.0%; Pred. No. 2.2e-83;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKFLIFAFPGVHLLSLCSGKAIKNGISKRTPEEIKKEELASCGDVAKAIINLAIVGKAQ 60
 DB 1 MKFLIFAFPGVHLLSLCSGKAIKNGISKRTPEEIKKEELASCGDVAKAIINLAIVGKAQ 60
 QY 61 NRSYERLALVDYVGPRLSSKNLEKAIQIMYQNLQDGLKYLEPVRIPHERGESB 120
 DB 61 NRSYERLALVDYVGPRLSSKNLEKAIQIMYQNLQDGLKYLEPVRIPHERGESB 120
 QY 121 VMLEPRHKAIALIGLSSIGTPPEGITAEVLVVTSPDELORRASEARGKIVVYNO 175
 DB 121 VMLEPRHKAIALIGLSSIGTPPEGITAEVLVVTSPDELORRASEARGKIVVYNO 175
 DE 08-MAY-2000 (first entry)
 XX
 XX Human signal peptide containing protein HSP-31 SEQ ID NO:31.
 KW Human; signal peptide-containing protein; HSP; diagnosis; cancer;
 KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;
 KW antimicrobial; neuroprotective; cardiovascular; hepatotropic;
 KW antiaesthetic; gene therapy; cell proliferation; neurological disorder;
 KW reproductive disorder; developmental disorder; arteriosclerosis;
 KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
 KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
 KW Parkinson's disease; Huntington's disease; ovulatory defect;
 KW muscular dystrophy.
 KM
 XX
 OS Homo sapiens.
 XX
 PN WO200000610-A2.
 XX
 PD 06-JAN-2000.
 XX
 PF 25-JUN-1999; 99WO-US14484.
 XX
 PF 26-JUN-1998; 98US-0090762.
 XX
 PR 31-JUL-1998; 98US-0094983.

KW Chromosome 8q21.3; prostate cancer; neurodegenerative disease;
 KW Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
 KW peripheral neuropathy; Huntington's disease; acute brain injury;
 KW multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;
 KW therapy; diagnosis; nocitropic; neuroprotective; neuroleptic;
 KW antiparkinsonian; anticonvulsant; vasodilator.
 OS Homo sapiens.
 XX
 XX
 XX Key Location/Qualifiers
 XX Domain 3..24
 XX /note= "membrane-spanning domain"
 XX
 XX Modified-site 121
 XX /note= "N-glycosylated"
 XX
 XX Modified-site 179
 XX /note= "N-glycosylated"
 XX
 XX Modified-site 353
 XX /note= "N-glycosylated"
 XX
 XX Modified-site 356
 XX /note= "N-glycosylated"
 XX
 XX Modified-site 396
 XX /note= "N-glycosylated"
 XX
 XX WO200004157-A2.
 XX
 XX 27-JAN-2000.
 XX
 XX 14-JUL-1999; 99WO-GB02241.
 XX
 XX 14-JUL-1998; 98GB-0015284.
 XX
 XX (JANNC) JANSSEN PHARM NV.
 XX
 XX Pangalos M, Neefs JEFM, Peeters DCG.
 XX
 XX WPI: 2000-182424/16.
 XX
 XX N-PSDB; AA258313.
 XX
 XX New human N-acetylated alpha-linked acidic dipeptidases for treating
 XX PT neural disorders e.g. Alzheimer's disease, schizophrenia and
 XX PT Parkinson's disease
 XX
 XX Claim 15; Fig 5; 95pp; English.
 XX
 XX The present sequence is that predicted for human N-acetylated
 XX CC alpha-linked acidic dipeptidase IV (NALAD-ase IV) on the basis of
 XX CC isolated cDNA (see AA258313). NALAD-ase IV is predicted to be a
 XX CC type II integral membrane protein of mol.wt. 51.9 kDa and pI 5.99.
 XX CC Expression was low in all tissues examined by RT-PCR. The
 XX CC invention provides human NALAD-ase I, II and IV cDNAs and encoded
 XX CC polypeptides, as well as vectors, host cells, transgenic organisms,
 XX CC antisense nucleic acids, agonists and antagonists. These are useful
 XX CC for treating neural disorders such as Alzheimer's disease,
 XX CC schizophrenia, ALS, Parkinson's disease, peripheral neuropathy,
 XX CC Huntington's disease, acute brain injury, multiple sclerosis,
 XX CC exposure to neurotoxins, peripheral nerve trauma, ischaemia or
 XX CC dementia (claimed).
 XX
 XX
 XX Sequence 472 AA;
 XX
 XX Query Match 100.0%; Score 883; DB 21; Length 472;
 XX Best Local Similarity 100.0%; Pred. No. 2.2e-83;
 XX Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 MKFLIPAFGVHLLSLCSGKALICNGISKRTFEEIKELASCGDVAKIINLAVYGKQ 60
 XX 1 MKFLIPAFGVHLLSLCSGKALICNGISKRTFEEIKELASCGDVAKIINLAVYGKQ 60
 XX 1 NRSYERLALVDYVGPRLSSGSKNLEKATQIYQNLQDGLKRVHLEPRIMHMERGES 120
 XX 61 NRSYERLALVDYVGPRLSSGSKNLEKATQIYQNLQDGLKRVHLEPRIMHMERGES 120
 XX 61 NRSYERLALVDYVGPRLSSGSKNLEKATQIYQNLQDGLKRVHLEPRIMHMERGES 120
 XX 121 VMEPRILHKAIIAGLSSIGTPPEGITAEVLVWTSFDELQRRASERAGKIIVYVQ 175

DB 121 VMEPRILHKAIIAGLSSIGTPPEGITAEVLVWTSFDELQRRASERAGKIIVYVQ 175
 RESULT 6
 AA06058
 AA06058 standard; Protein; 472 AA.
 AC
 AC AA06058;
 DT 24-SEP-2001 (first entry)
 DE
 DE Human gene 18 encoded secreted protein HRACJ35, SEQ ID NO:120.
 XX
 XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KW cell culture; chemotaxis; vulnerrary; binding partner identification;
 KW gene therapy.
 XX
 XX Homo sapiens.
 XX
 XX OS
 XX
 XX Key Location/Qualifiers
 XX Peptide 1..24
 XX /label= signal_peptide
 XX Protein 25..472
 XX /label= Mature_human_secreted_protein
 XX
 XX WO200151504-A1.
 XX
 XX 19-JUL-2001.
 XX
 XX 12-JAN-2001; 2001WO-US00911.
 XX
 XX 13-JAN-2000; 2000US-0482273.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Endress GA, Muscenski M, Ebner R;
 XX
 XX WPI: 2001-425865/45.
 XX
 XX N-PSDB; AAD11647.
 XX
 XX Isolated nucleic acid molecule encoding a human secreted protein is
 XX PT used in preventing, treating or ameliorating a medical condition -
 XX
 XX Claim 11; Page 750-752; 864pp; English.
 XX
 XX AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
 XX CC protein genes, and AAD06041-AA06132 represent the proteins they encode.
 XX CC AAD06133-AA06205 represent human secreted protein fragments.
 XX CC The secreted proteins and their genes are useful for preventing, treating
 XX CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 XX CC Pathological conditions can be diagnosed by determining the amount of the
 XX CC new protein in a sample or by determining the presence of mutations in
 XX CC the new genes. Specific uses are described for each of the 71 genes,
 XX CC based on the tissues in which they are most highly expressed, and include
 XX CC developing products for the diagnosis or treatment of proliferative
 XX CC disorders, cancer, tumours, foetal and developmental abnormalities,
 XX CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 XX CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 XX CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 XX CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 XX CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 XX CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
 XX CC pregnancy-related disorders, endocrine disorders, and infections. The

CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, for supporting cell culture of primary tissues,
 CC to regenerate tissues, to identify their cognate ligands or binding
 CC partners, and in chemotaxis, and can be used as a food additive or
 CC preservative to modify storage properties. Antibodies specific for a
 CC protein of the invention can be used in alleviating symptoms associated
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
 CC The present sequence represents a human secreted protein of
 CC the invention.

XX Sequence 472 AA;

Query Match 100.0%; Score 883; DB 22; Length 472;

Best Local Similarity 100.0%; Pred. No. 2,2e-83; Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFPGVHLLSCGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLAVYKQAQ 60

Db 1 MKFLIFAFPGVHLLSCGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLAVYKQAQ 60

QY 61 NRSYERLALLVDYTGPRLSGSKNLEKAIQIMYQNLQODGLEKYLEPVRIPMERGESA 120

Db 61 NRSYERLALLVDYTGPRLSGSKNLEKAIQIMYQNLQODGLEKYLEPVRIPMERGESA 120

QY 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPDELORRASEANGKIIVTNO 175

Db 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPDELORRASEANGKIIVTNO 175

RESULT 7

ABP61800 ID ABP61800 standard; Protein; 472 AA.

AC ABP61800;

DT 04-OCT-2002 (first entry)

DE Human polypeptide SEQ ID NO 154.

XX Human; cytostatic; antirheumatic; antiarthritic; vulnerary; analgesic;
 KM antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
 KM neuroprotective; nootropic; osteoprotic; haemostatic; vasotropic;
 KM antitumor; fungicide; antidiabetic; antiashtmatic; antiallergic;
 KM immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KM cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KM stem cell; growth factor; nervous system disease; neuropathy;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KM osteoporosis; severe combined immunodeficiency; SCID; infection;
 KM multiple sclerosis; rheumatoid arthritis; gene therapy.

XX Homo sapiens.

PN US2002065394-A1.

PD 30-MAY-2002.

PF 22-DEC-2000; 2000US-0745763.

PR 18-MAR-1998; 98US-0040963.

XX (JACO/) JACOBS K.

PA (MCCO/) MCCOY J M.

PA (LAVA/) LAVALLIE E R.

PA (COLL/) COLLINS-RACIE L A.

PA (EVAN/) EVANS C.

PA (MERB/) MERBERG D.

PA (TREA/) TREACY M.

PA (SPAU/) SPAULDING V.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Spaulding V;

XX WPI; 2002-582343/62.

DR N-PSDB; ABQ92016.

PT Novel secreted or transmembrane protein and polynucleotide encoding the

PT protein, useful for diagnosis and treatment of neurological disorders,

PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis

PS Claim 51; Page 114-115; 284pp; English.

XX The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (cDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (I) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polypeptide of the invention.

SO Sequence 472 AA;

Query Match 100.0%; Score 883; DB 23; Length 472;

Best Local Similarity 100.0%; Pred. No. 2,2e-83; Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFPGVHLLSCGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLAVYKQAQ 60

Db 1 MKFLIFAFPGVHLLSCGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLAVYKQAQ 60

QY 61 NRSYERLALLVDYTGPRLSGSKNLEKAIQIMYQNLQODGLEKYLEPVRIPMERGESA 120

Db 61 NRSYERLALLVDYTGPRLSGSKNLEKAIQIMYQNLQODGLEKYLEPVRIPMERGESA 120

QY 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPDELORRASEANGKIIVTNO 175

Db 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVTSPDELORRASEANGKIIVTNO 175

RESULT 8

ABG33880 ID ABG33880 standard; Protein; 472 AA.

AC ABG33880;

DT 15-JUL-2002 (first entry)

DE Human secreted protein encoded by gene 18 #1.

XX Human; secreted protein; gene therapy; immunosuppressive;

KM antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
KM vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
KM virocidic; fungicide; ophthalmological; autoimmune disease; neoplasm;
KM rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KM cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KM angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KM ocular disorder; corneal infection; wound healing; skin aging;
KM epithelial cell proliferation; food additive.
XX
OS Homo sapiens.
XX
PN WO200226931-A2.
XX
PD 04-APR-2002.
XX
PP 24-SEP-2001; 2001WO-US29871.
XX
PR 25-SEP-2000; 2000US-234925P.
PR 12-JAN-2001; 2001WO-US00911.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Komatsu Jis G, Duan DR, Rosen CA, Moore PA, Shi Y,
PI Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR,
PI Endress GA, Mucenski M, Ebner R;
XX
DR WPI: 2002-362489/39.
DR N-PSDB; ABR69743.
XX
PT Novel 71 isolated secreted polypeptides and polynucleotides encoding
PT the polypeptides, useful for treating Huntington's disease, sepsis,
PT meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
PT aschma
XX
PS Claim 11, Page 1231-1232; 1478pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (or its
CC fragment, homologue complement or allelic variant) encoding a human
CC secreted protein (and its fragment, domain, epitope, variant, secreted
CC form and species variant). Also included are a recombinant vector
CC comprising the nucleic acid, a recombinant host cell comprising the
CC vector, an antibody against the secreted protein, a recombinant host cell
CC that expresses the secreted protein and a method of identifying a binding
CC partner of the secreted protein. The nucleic acid and protein are used to
CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
CC for example autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC ocular disorders e.g. corneal infection. Many other diseases and
CC disorders are listed in the specification. The polypeptides can also be
CC used to aid wound healing an epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities. The present
CC sequence represents a novel human secreted protein of the invention.
XX
SQ Sequence 472 AA;
Query Match 100.0%; Score 883; DB 23; Length 472;
Best Local Similarity 100.0%; Pred. No. 2.2e-83;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKFLFAFGGVLHSLSCGKAICNGSKRTFEEIKKEIASCCGVAKAIINLAVYGAQ 60
DB 1 MKFLFAFGGVLHSLSCGKAICNGSKRTFEEIKKEIASCCGVAKAIINLAVYGAQ 60
QY 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIWMYONLQDGLKXVHLEPVRIPIHMERGEESA 120
DB 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIWMYONLQDGLKXVHLEPVRIPIHMERGEESA 120

QY 121 VMLEPRHKAIIAGSSIGNPPEGITAEVLYVTSFPELORPASEARGKIVYVQ 175
DB 121 VMLEPRHKAIIAGSSIGNPPEGITAEVLYVTSFPELORPASEARGKIVYVQ 175
RESULT 9
ID ABR47892 standard; Protein; 472 AA.
XX
AC ABR47892;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted protein, SEQ ID 783.
XX
XX Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
KM vulnerary; antiinflammatory; nootropic; neuroprotective;
KM antiparkinsonian; gene therapy; human; cardiovascular disorder.
XX
OS Homo sapiens.
XX
PN WO200295010-A2.
XX
PD 28-NOV-2002.
XX
PP 19-MAR-2002; 2002WO-US09785.
XX
PR 21-MAR-2001; 2001US-277340P.
PR 19-JUL-2001; 2001US-306171P.
PR 13-NOV-2001; 2001US-331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM,
XX
DR WPI: 2003-129429/12.
XX
PT Novel human secreted proteins, useful for detecting, preventing,
PT diagnosing, prognosticating, treating and/or ameliorating
PT cardiovascular disorders such as arrhythmia -
XX
PS Claim 13; SEQ ID 783; 1881pp; English.
XX
CC The present invention relates to novel human secreted proteins
CC (ABR47633-ABR48145) and their coding sequences (ACC5034-ACC50856). The
CC proteins and their coding sequences are useful for the preparation of a
CC diagnostic or pharmaceutical composition for diagnosing or treating a
CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
CC coronary arteriosclerosis and myocardial ischaemia), neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, renal disorders,
CC proliferative disorders and/or cancerous diseases and conditions, for
CC wound healing and epithelial cell proliferation, to treat inflammation or
CC infection, for treating thrombosis and arteriosclerosis, for treating or
CC preventing neural damage which occurs in neuronal disorders or
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
CC disease, to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts, to prevent skin aging or hair loss, to
CC stimulate growth and differentiation of hematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism.
CC Note: The sequence data for this patent was published in electronic
CC format and is available from WIPD at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 472 AA;
Query Match 100.0%; Score 883; DB 24; Length 472;
Best Local Similarity 100.0%; Pred. No. 2.2e-83;

PR 15-JUL-1998; 98US-0092956.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y,
 PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE,
 PI Mucenski M, Endress GA, Soppet DR;
 XX
 DR WPI; 2000-161128/14.
 DR N-PSDB; AAZ98034.
 XX
 PT New isolated human genes, useful for diagnosis and treatment of, e.g.
 PT cancers, neurological or blood disorders
 XX
 PS Claim 11; Page 383-385; 494pp; English.
 XX
 CC The polynucleotide sequences given in AAZ98017 to AAZ98108 encode the
 CC human secreted proteins given in AAY87064 to AAY87223. Human secreted
 CC protein can have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;
 CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
 CC antiallergic. The polynucleotides and their corresponding secreted
 CC polypeptides are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Human secreted protein s and their polynucleotides can
 CC be used for developing products for the diagnosis or treatment of cancer,
 CC tumours, neurodegenerative disorders, developmental abnormalities and
 CC foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation,
 CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The
 CC proteins or polynucleotides can also be used as food additives or
 CC preservatives. The proteins are also useful for identifying their
 CC binding partners. AAZ98008 to AAZ98016 and AAY87063 are sequence used in
 CC the exemplification of the present invention.
 XX
 SQ Sequence 473 AA;
 XX
 Query Match 100.0%; Score 883; DB 21; Length 473;
 Best Local Similarity 100.0%; Pred. No. 2.2e-83;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTFEEIKERIASCGDVAKAIINLAVYGAQ 60
 DB 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTFEEIKERIASCGDVAKAIINLAVYGAQ 60
 QY 61 NRSYERLALLVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKHLVPRIPHMERGES 120
 DB 61 NRSYERLALLVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKHLVPRIPHMERGES 120
 QY 121 VMLPRIRHKTAIILGLSSIGTPPEGITAEVLVWTSFDELQRRASEAKGIWVYNQ 175
 DB 121 VMLPRIRHKTAIILGLSSIGTPPEGITAEVLVWTSFDELQRRASEAKGIWVYNQ 175
 RESULT 14
 ABG09405
 ID ABG09405 standard; Protein; 474 AA.
 AC- ABG09405;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #9396.
 XX
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX

OS Homo sapiens.
 XX
 EN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS73592.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID NO 39764; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations in
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 474 AA;
 XX
 Query Match 100.0%; Score 883; DB 22; Length 474;
 Best Local Similarity 100.0%; Pred. No. 2.2e-83;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTFEEIKERIASCGDVAKAIINLAVYGAQ 60
 DB 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTFEEIKERIASCGDVAKAIINLAVYGAQ 60
 QY 61 NRSYERLALLVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKHLVPRIPHMERGES 120
 DB 61 NRSYERLALLVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKHLVPRIPHMERGES 120
 QY 121 VMLPRIRHKTAIILGLSSIGTPPEGITAEVLVWTSFDELQRRASEAKGIWVYNQ 175
 DB 121 VMLPRIRHKTAIILGLSSIGTPPEGITAEVLVWTSFDELQRRASEAKGIWVYNQ 175
 RESULT 15
 AAM25762
 ID AAM25762 standard; Protein; 481 AA.
 AC AAM25762;
 XX
 DT 16-OCT-2001 (first entry)
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 22, 2003, 11:37:50 ; Search time 14.3354 Seconds
(without alignments)
1173.984 Million cell updates/sec

Title: US-09-745-763-copy_1_175
Perfect score: 883
Sequence: 1 MKFLIFAFGFGVHLISLCSG.....FDEIQRRASEARKIVVYQ 175

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_76:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	217	24.6	493	2	T46974
2	163.5	18.6	467	2	G87564
3	93.5	10.6	557	2	G72356
4	92.5	10.5	1295	2	T24587
5	92	10.4	920	2	F81725
6	89.5	10.1	215	2	G83658
7	87	9.9	339	2	S27738
8	85.5	9.7	388	2	A80103
9	84.5	9.6	326	2	T36497
10	84.5	9.6	765	2	S74598
11	82.5	9.3	350	2	A98976
12	82.5	9.3	350	2	G85822
13	82	9.3	821	1	S76330
14	82	9.3	864	2	T48474
15	81.5	9.2	162	2	D64577
16	81	9.2	809	2	G86603
17	81	9.2	809	2	D72022
18	81	9.2	1132	2	D88250
19	81	9.2	1170	2	T25043
20	80	9.1	1486	2	C64832
21	80	9.1	1486	2	G90754
22	80	9.1	1486	2	B85618
23	79	8.9	428	1	D71155
24	79	8.9	428	2	F69001
25	79	8.9	1488	2	A80615
26	78.5	8.9	284	2	F70024
27	78.5	8.9	369	1	FOADM2
28	78.5	8.9	370	2	H72748
29	78.5	8.9	1863	1	A58881

30	78	8.8	672	2	G87386	TPR domain protein
31	78	8.8	1812	2	I49350	breast/ovarian can
32	77.5	8.8	368	1	FOADM5	minor core protein
33	76.5	8.7	492	2	T26936	hypothetical prote
34	76.5	8.7	761	2	E82167	5-methyltetrahydro
35	76.5	8.7	845	1	S34027	replication licens
36	76	8.6	605	2	D81748	DNA gyrase, chain
37	76	8.6	1535	2	T49042	hypothetical prote
38	75.5	8.6	414	2	F64386	hypothetical prote
39	75.5	8.6	579	2	S37953	hypothetical prote
40	75.5	8.6	730	2	G75292	excinuclease ABC
41	75	8.5	180	2	S39775	alpha-s2-casein fo
42	75	8.5	227	2	D45522	variant surface gl
43	75	8.5	250	2	T05100	ribosomal protein
44	75	8.5	450	2	B70318	chaperone Hsp10 - A
45	75	8.5	564	2	A34325	juvenile-hormone e

ALIGNMENTS

RESULT 1
T46974
leucyl aminopeptidase [EC 3.4.11.-] ES-62 precursor [validated] - nematode (Acanthocheil)
C:Species: Acanthocheilonema viteae
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 01-Feb-2002
C:Accession: T46974
R:Harnett, W.; Houston, K.M.; Tate, E.; Garate, T.; Apfel, H.; Adam, R.; Haslam, S.M.; P.
Mol. Biochem. Parasitol. 104, 11-23, 1999
A:Title: Molecular cloning and demonstration of an aminopeptidase activity in a filarial
A:Reference number: 224335; MUID:20055923; PMID:10589978
A:Accession: T46974
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-493 <HAR>
A:Cross-references: EMBL:AF077194; NID:G3386477; PIDN:AAC28365.1; PID:G3386478
C:Function:
A:Description: EC 3.4.11.- leucyl aminopeptidase [validated, MUID:20055923]
C:Keywords: aminopeptidase

Query Match 24.6%; Score 217; DB 2; Length 493;
Best Local Similarity 32.5%; Pred. No. 2.7e-11;
Matches 55; Conservative 32; Mismatches 58; Indels 24; Gaps 5;

Qy	15	LSLCSGKAIC-----KNGISKRTFEETKEETIASCGDVAKIINIAVGGKAQNRSTERYL 67
Db	14	LTVALGAALVDPDKTAPAKNYIQETFGKVAR-----LIQYITKGEVGLAYQWL 62
Qy	68	ALAVDTVGPRLSGSKNLEKAIQIMWYQLODGLKLVLPV-RIPMERGEESAVMLEPR 126
Db	63	SKLVDFGCHMVGSDSLKSIAPLESBKNDNPKVTEVPNLPVHVKGNDVEMTEPR 122
Qy	127	IKHTAIIIGLSSISGTPEEGITAEVLVYTSFDEIQRRASEARKIVVYQ 175
Db	123	NQRLNVALIG---GSEBASATGEVTVYLDLVV-KDDVVRGKIVVYQAQ 166

RESULT 2
G87564
aminopeptidase, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: G87564
R:Niernern, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.H.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon.
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapito, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: G87564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <STO>

A:Cross-references: GB:AE005673; NID:g13424107; PIDN:AAK2451.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2544

Query Match 18.5%; Score 163.5; DB 2; Length 467;
Best Local Similarity 35.5%; Pred. No. 1.3e-06;
Matches 38; Conservative 17; Mismatches 51; Indels 1; Gaps 1;

QY 70 LVDVGPRLSGSKNLEKAIQIMYQNLQODGLEKYLEPRIPMERGESAVMLEPRIRHK 129
DB 49 LTTNIGPRLVSPAPAKKMSVAKFKALGFTNIVDEPAKFSWRSGESALVAPVAMK 108

QY 130 IAIIGLSSIGTPEEGITAEVLVWTSFDELQRASEA-RGKIVVNO 175
DB 109 LGAVGLGRTVTRPAGIGAEVALFKTFADMAAPDGLKGIIVITQ 155

RESULT 3
G72356
penicillin-binding protein 2 - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: G72356
R.Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: G72356
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-557 <AB>
A:Cross-references: GB:AE001734; GB:AE000512; NID:g4981107; PIDN:AAJ5675.1; PID:g498110
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0590
C:Superfamily: penicillin-binding protein 3

Query Match 10.6%; Score 93.5; DB 2; Length 557;
Best Local Similarity 26.5%; Pred. No. 2.2;
Matches 48; Conservative 26; Mismatches 56; Indels 51; Gaps 11;

QY 20 GKAIKNGIS-----KRTPEIK-----EIASCGDVAKAIINLA---VYGA----- 59
DB 58 GKVLAKDEVVYVLDWNLISIDELKKTGCTPEEIIISLVKGBSTIVDKARADVISKAKRV 117

QY 60 ---QNRSYERLALLVDVGPRLSGSKNLEKAIQIMYQNLQODGLEKYLEP--- 107
DB 118 VMDYRKTEPLA-----PHVGVYNADRVGVGVSEVDEFLGTGDKVAVPEPSGA 170

QY 106 -----VRIPMERGESAVMLEPRIRHKAIIGLSSIGTPEEGT-----TAEVLVWTSFD 157
DB 171 ISSEVLRSPP-KPGEDVLTITDRIQKVAESL-EKIGNPGSVILSDVTRGIIALASFP 228

QY 158 E 158
DB 229 E 229

RESULT 4
T24587
hypothetical protein T06E4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24587
R.Lloyd, C.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19910
A:Accession: T24587
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1295 <WIL>

A:Cross-references: EMBL:Z70756; PIDN:CAA94789.1; GSPDB:GN00023; CESP:T06E4.1
A:Experimental source: clone T06E4
C:Genetics:
A:Gene: CESP:T06E4.1
A:Map position: 5
A:Introns: 17/3; 218/2; 385/3; 481/3; 946/3; 1034/3; 1222/2

Query Match 10.5%; Score 92.5; DB 2; Length 1295;
Best Local Similarity 20.4%; Pred. No. 7.8;
Matches 33; Conservative 44; Mismatches 54; Indels 31; Gaps 5;

QY 23 ICKNGISKRTPEEIKIEIASCGDVAKAIINLAVYKAQNSYERLALLVDVGPRLSGSK 82
DB 774 VMESEVLKESPEALQLEISASQEVRSVDAVOEK-----DGLRLVDTLKIKI---E 824

QY 83 NLEKAIQIMYQNLQODGLEKYLEPRIPMERGESAVMLEPRIRHKAIIGLSSIG 139
DB 825 DTERKA-----QDLQSSVEIKQLQDLQNFQKQAEVLESINERTLSSHKRDWVALASQL 880

QY 140 -----GTPPEGITAEVLVWTSFDELQRASEA 166
DB 881 EELQKLVGSGOVENYKSEILGKIKMKWVDLAKLGA 922

RESULT 5
P81725
cell division protein FtsH, probable TC0229 [imported] - Chlamydia muridarum (strain Nigg
P81725
C:Species: Chlamydia muridarum, Chlamydia trachomatis Morn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C:Accession: P81725
R.Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, I
C.; Dodson, R.; Gwin, M.; Nelson, M.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis Morn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Cross-references: GB:AE002290; GB:AE002160; NID:g7190267; PIDN:AAJ39101.1; PID:g719026
A:Experimental source: strain Nigg (Morn)
C:Genetics:
A:Residues: 1-920 <TE>
A:Molecule type: DNA
A:Status: preliminary
A:Accession: P81725
A:Cross-references: GB:AE002290; GB:AE002160; NID:g7190267; PIDN:AAJ39101.1; PID:g719026
A:Experimental source: strain Nigg (Morn)
C:Genetics:
A:Gene: TC0229

Query Match 10.4%; Score 92; DB 2; Length 920;
Best Local Similarity 24.2%; Pred. No. 5.6;
Matches 46; Conservative 36; Mismatches 58; Indels 50; Gaps 9;

QY 16 SLCSGKAICKNG--ISKRTPE--EIKIEIASCGDVAKAIINLAVYG--KAQNSYERL- 67
DB 138 SAISGMIPESGYTISRITDGLSVLEPIVHGPSVQIVNLATLEDRVNSLPSASLSR 197

QY 68 -----ALLVDVGPRLS-GSKNLEKAIQIMYQNLQODGLEKYLEPRIPMERGES 119
DB 198 VFGSDYALIGKYLSPALGIGSESIKKEIKDLYQOVESLTVQV-----EGQA 246

QY 120 AVMLE---PRHKAIIGLSSIGTPEEGITAEVL-----VWTSFDELQR 161
DB 247 ITLYKYVLETTLRIISL-----SLVSPEDGARFNQLRSVRLYREDFNRCMKILGSDSIOA 301

QY 162 RASEARGKIV 171
DB 302 QLDKIRGEIV 311

RESULT 6
C83658
general stress protein ctc [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Dec-2002
C:Accession: C83658
R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: C83658
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-215 <STO>
 A:Cross-references: GB:AB001507; GB:BA000004; NID:g10172612; PIDN:BA03786.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: ctc
 C:Superfamily: 50S ribosomal protein L25

Query Match 10.1%; Score 89.5; DB 2; Length 215;
 Best Local Similarity 28.0%; Pred. No. 1.5;
 Matches 44; Conservative 23; Mismatches 57; Indels 33; Gaps 8;

QY 44 GDVAKAI-----INLAVG---KAONRYERLALL--VDYTPRLSGSKNLEKAI--QIM 91
 DB 15 GSIVTRKIRKQGYVAVAVGKTKSOPISVEAVDFLKTREVGRNGLISLEVEKGTQKQVM 74
 QY 92 YQNIQODGL--EKVHLE-----PVRIPHWERGEASAVMLEPRIHKIALIIGL 135
 DB 75 VHDIQMPLKADYHIDPEFVDMSSSEIRANVPVRLTGARVSGCVLSQLMETITVRS 134
 QY 136 GSSIGTPPEGITAEVLVVTSPFDELQRRASERAKIV 172
 DB 135 PADI---PEBITLVSSLAIGDSIQIR--DVRGNV 166

RESULT 7
 S27738
 hypothetical protein D - *Thermus aquaticus*
 C:Species: *Thermus aquaticus*
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Jun-1993
 C:Accession: S27738
 R:Ashby, M.K.; Bergquist, P.L.
 submitted to the EMBL Data Library, July 1990
 A:Description: Cloning and sequence of IS1000 from *Thermus thermophilus* HB8.
 A:Reference number: S27735
 A:Accession: S27738
 A:Molecule type: DNA
 A:Residues: 1-339 <ASH>
 A:Cross-references: EMBL:M33159
 A:Note: the source is designated as *Thermus thermophilus*
 C:Genetics:
 A:Start codon: CTG

Query Match 9.9%; Score 87; DB 2; Length 339;
 Best Local Similarity 26.5%; Pred. No. 4.5;
 Matches 40; Conservative 21; Mismatches 52; Indels 38; Gaps 6;

QY 10 GCVHLISCSGKAIKNGISRTFEEIKEL-----ASGDVAKAI--INLAVG 57
 DB 6 GGLFLSLIMEE-----RTLEEVEVFKRGHGLVPKKELEDLAKALEAGLLTE 55
 QY 58 KAONRYERLALLVDYTPRLSGSKNLEKAIQIMYQNIQODGLEKVNLEPRIPHWERGE 117
 DB 56 KVEARLKEEKKLKRERWRLAG-----LSTPEGEREA--RAFLAFTASTPGE 104
 QY 118 ESAVMLEP-----RIHKIALIIGLSSIGTP 143
 DB 105 EARVLMPHLEPRSNVPEYGAALALEKTPP 135

RESULT 8
 AB0103
 probable acetylglucosamine-6-phosphate deacetylase YPO0838 [imported] - *Yersinia pestis*
 C:Species: *Yersinia pestis*
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
 C:Accession: AB0103
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.
 demo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,

Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AB0103
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-388 <KIR>
 A:Cross-references: GB:AL590842; PIDN:CAC89685.1; PID:g15978912; GSPDB:GN00175
 C:Genetics:
 A:Gene: YPO0838

Query Match 9.7%; Score 85.5; DB 2; Length 388;
 Best Local Similarity 24.6%; Pred. No. 7.2;
 Matches 43; Conservative 26; Mismatches 53; Indels 53; Gaps 7;

QY 11 GCVHLISCSGKAIKNGISRTF--EIKELISCGVAAIINLAVYKAONRYERLA 68
 DB 218 GVHLYNGMSGHLHREPCCGAVLYHDLAELADGIVHVMNVLAV---RMKGYRIA 273
 QY 69 ILVD-----TVGPRLSGSK-NLEKAIQIMYQNIQODGL 100
 DB 274 LITDQKAGGLGGRYILGKHHTVRCGEARTDDGSLAGSTCSLDALRMVIGHAQ---- 329
 QY 101 EKVHLEPRIPHWERGEASAVMLEPRIHKIALIIGLSSIGTPPEGITAEVLVVT 155
 DB 330 -----VPEWE-----AVQMSAV--PAAYIGLASTLISIQGAQASVMVMS 369

RESULT 9
 T36497
 probable branched-chain alpha keto acid dehydrogenase E1 beta chain - *Streptomyces coelic*
 C:Species: *Streptomyces coelicolor*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T36497
 R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1999
 A:Reference number: Z21608
 A:Accession: T36497
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-326 <SAU>
 A:Cross-references: EMBL:AL096822; PIDN:CAB46939.1; GSPDB:GN00070; SCODEB:SCGD3.17C
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: bkdB1, SCODEB:SCGD3.17C
 C:Superfamily: pyruvate dehydrogenase (lipoamide) beta chain

Query Match 9.6%; Score 84.5; DB 2; Length 326;
 Best Local Similarity 20.9%; Pred. No. 7.1;
 Matches 36; Conservative 30; Mismatches 73; Indels 33; Gaps 6;

QY 33 FEEIKELIA-----SCGDVAKAIINLAVY---GKAONRYERLALLVDYTPRLSGSKN 83
 DB 93 YDQIVTQLADQKARSISLCKVMPVIRIPYGGGIGAVHSHSEPEALPAHVAAGLKVSPSN 152
 QY 84 LEXAIQIMYQNIQODGLEKVNLEPRIPHWERGEASAVMLEPRIHKIALIIGLSSIG 139
 DB 153 AADAYMMQQAIGSDD--PVITYFEPR--RYWDKAEVDKELAPGPLHTARVREEGDTLLAA 210
 QY 140 -----GTPPEGITAEVLVVTSPFDELQRRASERAKIV 173
 DB 211 YGPVVKLCREVAADAAEGRSLEVDLRSISPDFTIQASVETRLIYVH 262

RESULT 10
 S74598
 hypothetical protein s11040 - *Synechocystis* sp. (strain PCC 6803)
 C:Species: *Synechocystis* sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S74598
 R:Kaneko, T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 o, K.; Okumura, S.; Shimpou, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
A:Accession: S74322; MUID:97061201; PMID:8905231
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-765 <KAN>
A:Cross-references: EMBL:D90900; GB:AB001339; NID:g1651768; PIDN:BA16750.1; PID:d101748
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 9.3%; Score 82.5; DB 2; Length 765;
Best Local Similarity 20.6%; Pred. No. 12;
Matches 45; Conservative 38; Mismatches 64; Indels 71; Gaps 11;

QY 4 LIFAFGGVHLLSCGKAIKNGISKR-----TFEIKERIASCGDVAKAIINLAVY--- 56
DB 548 LVLASFQGVAFVLAFLGRVVDMLNGTLIMTDRYAIGDVQIDVGLVENMNIYITQ 607
QY 57 -----GK---AQN---RSYERLALVDVTVGPRLSGSKNLEKA---IOIMYQNL 95
DB 608 LRPEGRISTTPNGKISVQVNLTKOMSHAEFTV-----IDQSNVDKALMLIRVSEOM 662
QY 96 QODGL-EKVHLEPVRI-----PHERGEESAVMLEPRIHKA 131
DB 663 REDPLWQKRI-LEPAALIGVDIDASKGIRLQVWIKTQAGQHPVGRERIRLM-----KKA 716
QY 132 ILGSSIGTPPEGITAEVLVVTSPDELORASEARKK 169
DB 717 FELAGIALGAPQORIV-----YHHGQKSHSNGK 746

RESULT 11
A:Accession: A98976
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
A:Reference number: A98629; MUID:21156231; PMID:11258796
A:Accession: A98976
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA36200.1; PID:g13362245; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
A:Genetics: C:Genetics: A:Gene: EC62777

Query Match 9.3%; Score 82.5; DB 2; Length 350;
Best Local Similarity 23.0%; Pred. No. 12;
Matches 35; Conservative 24; Mismatches 32; Indels 61; Gaps 6;

QY 28 ISKTFEIKERIASCGDVAKAIINLAVYGAQNSYERLALVDVTVGPRLSGSKNLEKA 87
DB 190 LSEVDLDEVEREMSLG---SYIALNLEGGSDQRTFS-----LSIAENL--- 230
QY 88 IOIMYQNLQODGLEKVHLE---PVRIIPHERGEESA-----VMLEPRIHKA- 131
DB 231 -----IKISETDIPVIYVHGPKGEDAKARLVDCNNVYRLSLSPSKRSA 278
QY 132 -----ILGSSIGTPPEGITAE 149
DB 275 IIKDAVIAITPDTSTILHMASAVNTPVAIVAD-310

RESULT 12
G85822
A:Title: hypochlorite protein Z3137 [imported] - *Escherichia coli* (strain O157:H7, substrain EDLg
A:Species: *Escherichia coli*

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
A:Accession: G85822
A:Status: N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Hiller, D.; Grobeck, B.O.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85822
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <STO>
A:Cross-references: GB:AB005174; NID:g12516156; PIDN:AAG57043.1; GSPDB:GN00145; UMG:P.Z31.
A:Experimental source: strain O157:H7, substrain EDL933
A:Genetics: C:Genetics: A:Gene: Z3137

Query Match 9.3%; Score 82.5; DB 2; Length 350;
Best Local Similarity 23.0%; Pred. No. 12;
Matches 35; Conservative 24; Mismatches 32; Indels 61; Gaps 6;

QY 28 ISKTFEIKERIASCGDVAKAIINLAVYGAQNSYERLALVDVTVGPRLSGSKNLEKA 87
DB 190 LSEVDLDEVEREMSLG---SYIALNLEGGSDQRTFS-----LSIAENL--- 230
QY 88 IOIMYQNLQODGLEKVHLE---PVRIIPHERGEESA-----VMLEPRIHKA- 131
DB 231 -----IKISETDIPVIYVHGPKGEDAKARLVDCNNVYRLSLSPSKRSA 278
QY 132 -----ILGSSIGTPPEGITAE 149
DB 279 IIKDAVIAITPDTSTILHMASAVNTPVAIVAD 310

RESULT 13
S76330
A:Accession: S76330
A:Title: endopeptidase Clp (EC 3.4.21.-) ATP-binding chain C [similarity] - *Synechocystis* sp. (str
N:Alternate names: ATP-dependent Clp proteinase regulatory chain
N:Contains: adenosinetriphosphatase (EC 3.6.1.3)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 19-Jan-2001
A:Accession: S76330
A:Status: N.T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
R.Kaneko, T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
C.K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76330
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-821 <KAN>
A:Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BA1182.1; PID:g100155
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
A:Function: C:Function: A:Function: allows clp to hydrolyze polypeptides and proteins, probably by a chaperon
e activity. ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller
C:Superfamily: endopeptidase Clp ATP-binding chain
C:Keywords: ATP; hydrolase; molecular chaperone; nucleotide binding; P-loop; serine prot
F:208-215/Region: nucleotide-binding motif A (P-loop)
F:275-280/Region: nucleotide-binding motif B
F:545-552/Region: nucleotide-binding motif A (P-loop)
F:613-618/Region: nucleotide-binding motif B
F:714/Binding site: ATP (Lys) #status predicted
F:751/Binding site: ATP (Lys) #status predicted

Query Match 9.3%; Score 82; DB 1; Length 821;
Best Local Similarity 21.3%; Pred. No. 37;
Matches 43; Conservative 27; Mismatches 70; Indels 62; Gaps 6;

QY 19 SKAICNGISKTFEIKERIASCGDVAKAIINLAVYGAQNSYERLALVDVTVGPRL 78
DB 459 SKKAGNGEPEVTSSEIAIHVSWTCV-----VNLTSESESKLHMETLHQRLL 510

```

Db      20  IGSFTKENIK-EYAKC-----FFMSLSMFLBEOEKNOOEFLBQDT-----KENOEEL 66

Qy      88  IQMNYNOODGLEKXHLPEVRIPHMEGESSAWLBERHUKIA-----LGSSICTPP 143
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      67  I-----KNIDTSIAKNOELEKISFEKXENKIDERYL--PKLRIVTHKLOBSITSIINTOL 120

Qy      144  EGITAEVLVVTSPDELCORRASEARSKIYV 172
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      121  ESFKDELDLSSVFEIQKRNTOIAVRLAI 149

Search completed: December 22, 2003, 11:53:01

```

06 28 TSKPTFEETKEETASCCDVAKATINI AVYCKRONPSYFRIALYDTVCBPISSCINU EKA 07

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: December 22, 2003, 11:34:20 ; Search time 11.0896 Seconds
(without alignments)
742.104 Million cell updates/sec

Title: US-09-745-763-COPY_1_175
Perfect score: 883
Sequence: 1 MKFLIFAFGCVHLLSLCSG.....FDELORRASEARKIVVYVQ 175

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.5	9.6	552	VNS1_EHDV2	P27585 epizootic h
2	81	9.2	809	MUDD_CHLPN	O94701 c murc/dcl
3	80	9.1	1486	MUKB_ECOLI	P25523 escherichia
4	78.5	8.9	284	DAF7_BACSU	O32114 bacillus su
5	78.5	8.9	369	VCOM_ADEB2	P03267 human adeno
6	78.5	8.9	370	EGSA_AERPE	O94722 aeropyrum p
7	78.5	8.9	1863	BRCI_HUMAN	P38398 homo sapien
8	78	8.8	1812	BRCI_MOUSE	P48754 mus musculu
9	77.5	8.8	368	VCOM_ADEB5	P24938 human adeno
10	76.5	8.7	738	AB10_HUMAN	O94718 homo sapien
11	76.5	8.7	761	METE_VIRCH	O94718 vibrio chol
12	76.5	8.7	845	CC47_YEAST	P38132 saccharomyc
13	75.5	8.6	414	Y694_METUA	O58105 methanococ
14	75.5	8.6	579	Y694_YEAST	P33343 saccharomyc
15	75.5	8.6	1286	SMC4_MOUSE	O86947 mus musculu
16	75	8.5	180	CAS2_RABIT	P50418 corycolagus
17	75	8.5	450	HSIU_AQUNE	O65574 aquifex aeo
18	75	8.5	807	STL_BACSU	P36430 bacillus su
19	75	8.5	804	MCN3_XENLA	P47339 xenopus lae
20	75	8.5	1510	MUKB_HAEIN	P45187 haemophilus
21	74.5	8.4	183	PYRR_DEIRA	O94709 deinoococcus
22	74.5	8.4	1967	VG50_YEAST	P53327 saccharomyc
23	74	8.4	355	KUC3_HUMAN	O94860 homo sapien
24	74	8.4	390	SCC1_HUMAN	P29508 homo sapien
25	73.5	8.3	234	RS2_YEAST	P81289 bacillus st
26	73	8.3	255	SOL4_YEAST	P53315 saccharomyc
27	73	8.3	344	Y613_METUA	O58030 methanococ
28	73	8.3	428	CLPX_STRCO	O94316 streptomyc
29	73	8.3	476	MRPB_NEURC	P11913 neurospora
30	73	8.3	670	CKX1_ARATH	O30054 arabidopsis
31	73	8.3	893	MVP_HUMAN	O14764 homo sapien
32	72.5	8.2	205	RIVA_MYCBE	O49424 mycoplasma
33	72.5	8.2	351	DPO4_ECOS7	O84791 escherichia

34	72.5	8.2	351	1	DPO4_ECOL6	P59477 escherichia
35	72.5	8.2	441	1	ARGD_SCHPO	O74548 schizosacch
36	72.5	8.2	510	1	PURI_CHICK	P28173 gallus gall
37	72.5	8.2	881	1	YEET_HAEIN	P44288 haemophilus
38	72	8.2	249	1	RS6_ONCMY	O94752 oncorhynch
39	72	8.2	429	1	YU64_AQUNE	O67776 aquifex aeo
40	72	8.2	886	1	ODPI_ECOLI	P06958 escherichia
41	71.5	8.1	223	1	GIDE_CORG	O8153 corynebacte
42	71.5	8.1	303	1	CHEV_BACSU	P37599 bacillus su
43	71.5	8.1	411	1	MANA_CAEEL	P34550 caenorhabd
44	71.5	8.1	513	1	CP12_MOUSE	P00186 mus musculu
45	71.5	8.1	1288	1	SMC4_HUMAN	O94718 homo sapien

ALIGNMENTS

RESULT 1	VNS1_EHDV2	STANDARD;	PRT;	552 AA.
ID	VNS1_EHDV2			
AC	P27585;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	01-DEC-1992 (Rel. 24, Last annotation update)			
DE	Nonstructural protein NS1 (Hydrophobic tubular protein).			
GN	S6.			
OS	Epizootic hemorrhagic disease virus (serotype 2 / strain Alberta)			
OS	(Ehdv-2).			
OC	Viruses; dsRNA viruses; Reoviridae; Orbivirus.			
OX	NCBI_TaxID=10910;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91253267; PubMed=1645906;			
RA	Nel L.H., Picard L.A., Huismans H.;			
RT	"A characterization of the nonstructural protein from which the			
RT	virus-specified tubules in epizootic haemorrhagic disease			
RT	virus-infected cells are composed."			
RL	Virus Res. 18:219-230(1991).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M69085; AAA43000.1; -			
DR	EMBL; X59000; CAA41746.1; -			
DR	InterPro: IPR002630; Orbi NS1.			
DR	Pfam: PF01718; Orbi NS1; I.			
KW	Nonstructural protein.			
SQ	SEQUENCE 552 AA; 64553 MW; 6D440708327676F CRC64;			
Query Match	9.6%; Score 84.5; DB 1; Length 552;			
Best local Similarity	25.0%; Pred. No. 4.1;			
Matches	20; Conservative 19; Mismatches 46; Indels 19; Gaps 4;			
Qy	8 FFGVHLLSCSKAIKNGISKRTEPEI--KEEIASCGDVAKAIINLAVYKAQRNSY 64			
Db	329 FMOGRHRCQCL---FLKNGCDRETFYHIDVTRSEIWCSTVQVMIG-----EH 374			
Qy	65 ERLALVDVTPGRPLSGSKNLEKAIQIMYQNLQDDGKHLDEVRIPHWERG 116			
Db	375 VDISLPQVKI--KLITGHEHGRASDHFKYNATFGMALIRTAIQIHWIRG 424			
RESULT 2	MUDD_CHLPN	STANDARD;	PRT;	809 AA.
ID	MUDD_CHLPN			
AC	O92701; Q9J0B5;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			

Query Match	9.2%	Score 81	DB 1	Length 809
Best Local Similarity	24.2%	Pred. No. 13		
Matches 48	Conservative 28	Mismatches 62	Indels 60	Gaps 9
QY	17	LCSGKAIKNGIKSRTETEEIKEELA-SCGDVAKAIINL-----AVYKQAQ	60	
DB	457	VCGGKS-CEDHISLSAQHWSKYISPEPYDVSYFIINROGLMRTGKDPHLIEFTQSDSP	515	
QY	61	NRSYERIALL-VDPVGGRLSGSKYLKKAQIMYONLQOD-----	98	
DB	516	LSSEIASALAKVDDLFPVLHGGPREDDTQGFEEILKPYAGSLSLAATAMDKLTKRI	575	
QY	99	---GLEKVLHPYRIPHWERGESSAV--MLE---PRIKIIAIIIGVSSIGTPPEGITA	148	
DB	576	ASAVGVVVPVYQPLNLCFWKRNPCLQIQLIETPSFPMIYKTA--HIGSSIG-----	625	
QY	149	EVLVVTSPDEIQRASFA 166		
DB	626	-IFLVKDEKEIQEKISEA 642		
RESULT 3				
MUKB_ECOLI				
ID	MUKB_ECOLI	STANDARD;	PRT;	1486 AA.
AC	P22523;	P71227; P77164; Q47398;		
DT	01-AUG-1991 (Rel. 19, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cell division protein mukB.			
GN	MUKB OR B0924.			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
OX	NCBI_TaxID=562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=K12 / W3110.			
RC	STRAIN=K12 / W3110.			
RX	MEDLINE=9114703; Pubmed=1989883;			
RA	Niki H., Jaffe A., Imamura R., Ogura T., Hiraga S.;			
RT	"The new gene mukB codes for a 177 kd protein with coiled-coil			
RT	domains involved in chromosome partitioning of E. coli.";			
RL	EMBO J. 10:183-193(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A. AND MUTANTS MUKB33 AND MUKB106.			
RX	MEDLINE=95080615; Pubmed=7988894;			
RA	Yamanaka K., Mitani T., Feng J., Ogura T., Niki H., Hiraga S.;			
RT	"Two mutant alleles of mukB, a gene essential for chromosome			
RT	partition in Escherichia coli.";			
RL	FEMS Microbiol. Lett. 123:27-31(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=K12 / MG1655;			
RC	MEDLINE=94426617; Pubmed=9278503;			
RX	Baltner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Siao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.";			
RL	Science 277:1453-1474(1997).			

[4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashiimoto K.,
 RA Kimura S., Kitegawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiuchi T.,
 RT "A 718-Kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 RN [5]
 RP SEQUENCE OF 1-44 FROM N.A.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=94232180; PubMed=7513784;
 RA Feng J., Yamanaka K., Niki H., Ogura T., Hiraga S.,
 RT "New killing system controlled by two genes located immediately
 RT upstream of the mukB gene in *Escherichia coli*.";
 RL Mol. Gen. Genet. 243:136-147(1994).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1-227.
 RX MEDLINE=20015369; PubMed=10545328;
 RA van den Ent F., Lockhart A., Kendrick-Jones J., Loewe J.,
 RT "Crystal structure of the N-terminal domain of MukB: a protein
 RT involved in chromosome partitioning.";
 RL Structure 7:1181-1187(1999).
 CC -1- FUNCTION: ESSENTIAL FOR CHROMOSOME PARTITIONING. IMPLICATED IN
 CC ATP-DEPENDENT CHROMOSOME PARTITIONING DURING CELL DIVISION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

FT STRAND 18 23
 FT HELIX 25 33
 FT HELIX 38 51
 FT TURN 55 57
 FT HELIX 78 81
 FT STRAND 80 80
 FT STRAND 85 93
 FT TURN 95 96
 FT STRAND 99 108
 FT TURN 112 114
 FT STRAND 116 124
 FT TURN 128 129
 FT HELIX 132 136
 FT STRAND 137 138
 FT STRAND 146 147
 FT HELIX 150 158
 FT TURN 159 159
 FT TURN 161 162
 FT STRAND 164 167
 FT HELIX 171 180
 FT TURN 181 182
 FT STRAND 184 184
 FT HELIX 191 205
 FT STRAND 210 220
 FT STRAND 221 221
 SQ SEQUENCE 1486 AA; 170229 MW; 38C7874BB78D6D6 CRC64;
 Query Match 9.1%; Score 80; DB 1; Length 1486;
 Best Local Similarity 27.0%; Pred. No. 34;
 Matches 34; Conservative 17; Mismatches 33; Indels 42; Gaps 7;
 QY 52 NLAVYGAQRNSYERLALLVD-----TVGPRLSGSKN-LEKAIQIMONLODGLKXV 103
 DB 1304 NRIITFSALAKLQRLNPQIDMGQRTPQTIGEBLLRYNLEMEVEV---NRSGD--- 1356
 QY 104 HLEPVRIPHERGESESAVMLEPRHKAIALGLSSISCTPEEGITAEVLYVTSFDELQRA 163
 DB 1357 -----WLRASG-----LSTGEALGT---GMSILVMVQVMSDEBRR- 1391
 QY 164 SEARGK 169
 DB 1392 --LRGK 1395
 RESULT 4
 ID DAPF_BACSU STANDARD; PRT; 284 AA.
 AC 032114;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Diaminopimelate epimerase (EC 5.1.1.7) (DAP epimerase).
 GN DAPF.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxId=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borrilliet S., Bruschel L., Brans A., Braun W., Brignell S.C., Bron S.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Emlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gilm S.Y., Glaeser F., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.-J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kaashara Y., Klearr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portelle D., Porwolik S., Prescott A.M.,
 RA Pireson E., Fujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Setkovich J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Tosto V., Uchiyama S., Vandembol M., Vannier F., Vassarelli A.,
 RA Viari A., Wambuit R., Wedler E., Wedler H., Wellenreger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zunsstein E., Yoshikawa H., Zanchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).
 CC -1- CATALYTIC ACTIVITY: L1-2,6-diaminoheptanedioate = meso-
 CC diaminoheptanedioate.
 CC -1- PATHWAY: Biosynthesis of lysine from aspartate semialdehyde; sixth
 CC step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the diaminoheptamate epimerase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z99120; CAB15207.1; -
 DR PIR: F70024; F70024.
 DR HSSP: P44859; 1BW2.
 DR Subtilist; BG14048; dapF.
 DR HAMAP: MF_00197; -; 1.
 DR InterPro: IPR001653; DAP epimerase.
 DR Pfam: PF01678; DAP epimerase; 2.
 DR TIGRFAMs: TIGR00652; DapF; 1.
 DR PROSITE: PS01326; DAP EPIMERASE; 1.
 KW Isomerase; Lysine biosynthesis; Complete proteome.
 FT ACT SITE 76 76 BY SIMILARITY.
 FT ACT SITE 226 226 BY SIMILARITY.
 SQ SEQUENCE 284 AA; 30871 MW; 531D90C174C3B6 CRC64;
 Query Match 8.9%; Score 78.5; DB 1; Length 284;
 Best Local Similarity 18.8%; Pred. No. 6.4;
 Matches 36; Conservative 31; Mismatches 85; Indels 39; Gaps 5;
 QY 5 IFAPFGVHLSTLSCGKAIKNGISKRTPEEKEIKIASCQGVAKIINLAVYGRKQNSY 64
 DB 41 VYTGISDGLILICPSD---QAPVKKRIFNNDGSGKCGNGLRCVARYAY---EHLGV 93
 QY 65 ERLALLVDTVGPRLSGSKNLEKAIQIMYQNLQODLEKVLHPVRIPIHMERGEESAVMLE 124
 DB 94 EETSLIET---LSGLVKAIVQVNGKVNVTVMGMBRLTKSLPLMDGEHEETIWT 149
 QY 125 PRIHKAILGLSSIGTP-----PEGITAEVLVVTSPD 157
 DB 150 MAFGEVELTGVAVSMGNPHIVPIADIHQAPLTTLGPVLEKDPREPPEGINEVFETVNEQ 209
 QY 158 ELQRASAEARG 168
 DB 210 ELHFRVWE-RG 219
 RESULT 5
 VCOM_ADR02
 ID_VCOM_ADR02 STANDARD; PRT; 369 AA.
 AC P03267;
 DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Minor core protein (Protein V).
 GN PV.
 OS Human adenovirus type 2.
 CC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 CC NCBI_TaxID=10515;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85054835; PubMed=6094534;
 RA Alestrom P., Ausjalevi G., Lager M., Yeh-Kai L., Pettersson U.;
 RT "Genes encoding the core proteins of adenovirus type 2.";
 CC J. Biol. Chem. 259:13980-13985(1984).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J01917; AAA92213.1; -
 DR PIR: A03837; FOADM2.
 DR InterPro: IPR005608; Adeno PV.
 DR Pfam: PF03910; Adeno PV; 1.
 KW Core protein; Late protein.
 SQ SEQUENCE 369 AA; 41721 MW; 577C9B645B5E7DE CRC64;
 Query Match 8.9%; Score 78.5; DB 1; Length 369;
 Best Local Similarity 25.4%; Pred. No. 8.7;
 Matches 47; Conservative 23; Mismatches 58; Indels 57; Gaps 9;
 QY 29 SKRTPEEKEIKIASCQGV---AKIINLAVYGRKQNSYERLALLVDTVGPRLS----- 79
 DB 90 STRTYKRVYDEVGDEDLLEQANERLGEFAYGK---RHKMLALPLDEGNPTPLKPVTL 146
 QY 80 -----GSKNLEKAIQIMYQNLQ---QDLEKVLHPVRIPIHMERGEES 119
 DB 147 QQVPLTAPSEKRGKLRBSGDLAPVQLVPRKQRLDEVLKRTVP-----GLEP 198
 QY 120 AVMLEPRIHKAILGLG-----SSIGTPEGITAEVLVVTSS--FDELQRAS 164
 DB 199 EVVRVP-IKQVA-PGLGVQTVDVQIPTSSITSATATEGNETQSPVAAVADAQVAAA 256
 QY 165 EARGK 169
 DB 257 AAASK 261
 RESULT 6
 BGSA_AERPE
 ID_BGSA_AERPE STANDARD; PRT; 370 AA.
 AC Q9YER2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceral-1-phosphate dehydrogenase [NAD(P)] (EC 1.1.1.261) (Sn-
 DE glycerol-1-phosphate dehydrogenase) (G-1-P dehydrogenase)
 DE (Enantiomeric glycerophosphate synthase).
 GN BGSA OR APE0519.
 OS Aeropyrum pernix.
 CC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
 CC Desulfurococcaceae; Aeropyrum.
 CC NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=KL;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatawara Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anket A., Kosugi H.,
 RA Hasegawa A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,

RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 CC -1- FUNCTION: Responsible for the formation of archaea-specific
 CC glycerophosphate backbone of phospholipids, G-1-P, from
 CC dihydroxyacetonephosphate (DHAP) (By similarity).
 CC -1- CATALYTIC ACTIVITY: Sn-glycerol-1-phosphate + NAD(P)(+) =
 CC glycerone phosphate + NAD(P)H.
 CC -1- PATHWAY: De novo phospholipid biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
 CC -1- SIMILARITY: BELONGS TO THE GLYCEROL-1-PHOSPHATE DEHYDROGENASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AP000059; BAA79484.1; -
 DR PIR; H72748; H72748.
 DR HAMAP; MF 00497; -; 1.
 DR InterPro; IPR002658; DHQ synthase.
 DR Pfam; PF01761; DHQ synthase, 1.
 KW Phospholipid biosynthesis; Oxidoreductase; NADP; Complete proteome.
 SQ SEQUENCE 370 AA; 39352 MW; 0782087EB9FCBF01 CRC64;
 Query March 8.9%; Score 78.5; DB 1; Length 370;
 Best Local Similarity 30.4%; Pred. No. 8.8;
 Matches 42; Conservative 18; Mismatches 47; Indels 31; Gaps 7;
 QY 55 YGKAQNRSYERLALVDVTPRLSGSKULEKAIQIMYQNIQDGLKHPVRIIPHE 114
 DB 46 VSGVAQGSY---LVVVS--GPTVS-SKYFER---LRASLEABGL-TVGKIRIDATVE 94
 QY 115 RGEESA-VMEPRHKAIALGLSSIGRP-----EITAEVLVVT 154
 DB 95 TAEVVAEALLESRLBVAAGLGGKSIDVAKYASRAGSVFVSIPTVASHDITSPFSLK 154
 QY 155 SFDELQRRASEARGKIV 172
 DB 155 GFDKPISRPAKAPRAII 172
 RESULT 7
 BRCA1_HUMAN STANDARD; PRT; 1863 AA.
 ID BRCA1_HUMAN
 AC P38398;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Breast cancer type 1 susceptibility protein.
 GN BRCA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT ARG-1775.
 RX MEDLINE=95025896; PubMed=754954;
 RA Miki Y., Swensen J., Shattuck-Bidens D., Futreal P.A., Harshman K.,
 RA Tavtigian S., Liu Q., Cochran C., Bennett L.M., Ding W., Bell R.,
 RA Rosenthal J., Husey C., Tran T., McClure M., Frye C., Hattler T.,
 RA Phelps R., Haugen-Strano A., Katcher H., Yakumo K., Gholami Z.,
 RA Shaffer D., Stone S., Bayer S., Wray C., Bogden R., Dayananth P.,
 RA Ward V., Tonin P., Narod S., Bristow P.K., Norris F.H., Helvering L.,
 RA Morrison P., Roestek P., Lai M., Barrett J.C., Lewis C., Neuhausen S.,
 RA Cannon-Albright L., Goddard D., Wiseman R., Kamb A., Skolnick M.H.;
 RT "A strong candidate for the breast and ovarian cancer susceptibility

RT gene BRCA1.";
 RL Science 266:66-71(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97092865; PubMed=8938427;
 RA Smith T.M., Lee M.K., Szabo C.I., Jerome N., McEuen M., Taylor M.,
 RA Hood L., King M.-C.;
 RT "Complete genomic sequence and analysis of 117 kb of human DNA
 RT containing the gene BRCA1.";
 RL Genome Res. 6:1029-1049(1996).
 RN [3]
 RP FUNCTION AS A E2-DEPENDENT UBIQUITIN LIGASE.
 RX MEDLINE=99432238; PubMed=10500182;
 RA Lorick K.L., Jensen J.P., Fang S., Ong A.M., Hatakeyama S.,
 RA Weissman A.M.;
 RT "RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent
 RT ubiquitination.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:11364-11369(1999).
 RN [4]
 RP IDENTIFICATION OF BRCA1 AS MEMBER OF BASC.
 RX MEDLINE=20245492; PubMed=10783165;
 RA Wang Y., Cortez D., Yazdi P., Neff N., Elledge S.J., Qin J.;
 RT "BASC, a super complex of BRCA1-associated proteins involved in the
 RT recognition and repair of aberrant DNA structures.";
 RL Genes Dev. 14:927-939(2000).
 RN [5]
 RP INTERACTION WITH SMC11.
 RX MEDLINE=21866464; PubMed=11877377;
 RA Yazdi P.T., Wang Y., Zhao S., Patel N., Lee E.Y.-H.P., Qin J.;
 RT "SMC1 is a downstream effector in the ATM/ATR branch of the human
 RT S-phase checkpoint.";
 RL Genes Dev. 16:571-582(2002).
 RN [6]
 RP PHOSPHORYLATION BY ATM, AND MUTAGENESIS OF SER-1387, SER-1423 AND
 RP SER-1524.
 RX MEDLINE=22170551; PubMed=12183412;
 RA Xu B., O'Donnell A.H., Kim S.-T., Kastan M.B.;
 RT "Phosphorylation of serine 1387 in Brca1 is specifically required for
 RT the Atm-mediated S-phase checkpoint after ionizing irradiation.";
 RL Cancer Res. 62:4586-4591(2002).
 RN [7]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=96400954; PubMed=8807330;
 RA Couch F.J., Weber B.L.;
 RT "Mutations and polymorphisms in the familial early-onset breast
 RT cancer (BRCA1) gene. Breast Cancer Information Core.";
 RL Hum. Mutat. 8:8-18(1996).
 RN [8]
 RP VARIANTS LEU-1637, GLU-1708 AND ARG-1775.
 RX MEDLINE=95025878; PubMed=7939630;
 RA Futreal P.A., Liu Q., Shattuck-Bidens D., Cochran C., Harshman K.,
 RA Tavtigian S., Bennett L.M., Haugen-Strano A., Swensen J., Miki Y.,
 RA Eddington K., McClure M., Frye C., Weaver-Felhaus J., Ding W.,
 RA Gholami Z., Soederqvist P., Terry L., Jhanwar S., Berchuk A.,
 RA Iglehart J.D., Marks J., Ballinger D.G., Barrett J.C., Skolnick M.H.,
 RA Kamb A., Wiseman R.;
 RT "BRCA1 mutations in primary breast and ovarian carcinomas.";
 RL Science 266:120-122(1994).
 RN [9]
 RP VARIANTS BC GLY-64, AND VARIANTS ALA-772; ASN-1040 AND GLY-1443.
 RX MEDLINE=95201806; PubMed=7894491;
 RA Castilla L.H., Couch F.J., Erdos M.R., Hoekins K.F., Calzone K.,
 RA Garber J.E., Boyd J., Lubin M.B., Deshano M.L., Brody L.C.,
 RA Collins F.S., Weber B.L.;
 RT "Mutations in the BRCA1 gene in families with early-onset breast and
 RT ovarian cancer.";
 RL Nat. Genet. 8:387-391(1994).
 RN [10]
 RP VARIANT BC GLY-61, AND VARIANTS ARG-356; GLY-1038; ASN-1040; ARG-1183
 RP AND GLY-1613.
 RX MEDLINE=95201808; PubMed=7894493;
 RA Friedman L.S., Ostermeyer E.A., Szabo C.I., Dowd P., Lynch E.D.,
 RA Rowell S.E., King M.-C.;

Query	Local	Similarity	Score	DB 1	Length	1863
Best	Local	Similarity	22.8%	Pred. No. 62		
Matches	44	Conservative	34	Mismatches	76	Indels 39; Gaps 11;
QY	5	IFAFPGVHLLSLCSGAA--ICNGKSGKRT-----FEETKEKIAS--CG---DVAKAI	50			
DB	42	IFCFKFCMLKLNQKKGGSCQCPCLCNDITDKSLDSTRFSQIYVERLKTICAFOLDTGLE	101			
QY	51	INLAVYGCAQNRYSYERL--ALAVDTVGPRLSGSKNL-----EKAIGIWMYONL--	95			
DB	102	ANSTNFAKKNSEBHLKDEVSIIQSGNYRNRARLLQSEPPNDSLOETSLSVGLSNGT	161			
QY	96	QODGLEKVLHPVPIP--HMERGESASVMLEPRLIKIALILGSGS--IGTPEEGITAEVL	152			
DB	162	VRIITRTQRIQRIQKQTSYVIELGSSS---EDTVNKATVYSGVGDDELLQIRPGGRDEI--	216			
QY	153	VTSFDELQRRASE	165			
DB	217	--SIDSACKKACE	227			
RESULT 8						
BRCA1_MOUSE						
ID	BRCA1_MOUSE	STANDARD;	PRT;	1812	AA.	
AC	P48754; Q60957; Q60983;					
DT	01-FEB-1996 (Rel. 33, Created)					
DT	01-NOV-1997 (Rel. 35, Last sequence update)					
DT	15-SEP-2003 (Rel. 42, Last annotation update)					
DE	Breast cancer type 1 susceptibility protein homolog.					
GN	BRCA1.					
OS	Mus musculus (Mouse).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX	NCBI_TaxId=10090;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=C57BL/6; TISSUE=Embryo;					
RX	MEDLINE=96177659; PubMed=8634697;					
RA	Abel K.J., Xy J., Yin G.Y., Lyons R.H., Weisler M.H., Weber B.L.,					
RT	"Mouse Brca1: localization sequence analysis and identification of					
RT	evolutionarily conserved domains."					

RL Hum. Mol. Genet. 4:2265-2273(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=96177660; PubMed=8634698;
 RA Sharan S.K., Wims M., Bradley A.;
 RT "Murine Brca1: sequence and significance for human missense
 mutations.";
 RL Hum. Mol. Genet. 4:2275-2278(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=96121367; PubMed=8575748;
 RA Bennett L.M., Haugen-Strano A., Cochran C., Brownlee H.A.,
 RA Fiedorek F.T. Jr., Wiseman R.W.;
 RT "Isolation of the mouse homologue of BRCA1 and genetic mapping to
 mouse chromosome 11.";
 RL Genomics 29:576-581(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ; TISSUE=Embryo;
 RX MEDLINE=96067162; PubMed=7590247;
 RA Lane T.F., Deng C., Elson A., Lyu M.S., Kozak C.A., Leder P.;
 RT "Expression of Brca1 is associated with terminal differentiation of
 ectodermally and mesodermally derived tissues in mice.";
 RL Genes Dev. 9:2712-2722(1995).
 RN [5]
 RP SEQUENCE OF 727-1111 FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Embryo;
 RX MEDLINE=96021028; PubMed=7550308;
 RA Marquis S.T., Rajan J.V., Wynshaw-Boris A., Xu J., Yin G.Y.,
 RA Abel K.J., Weber B.L., Chodosh L.A.;
 RT "The developmental pattern of Brca1 expression implies a role in
 RT differentiation of the breast and other tissues.";
 RL Nat. Genet. 11:17-26(1995).
 RN [6]
 RP SEQUENCE OF 789-1250 FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=96163506; PubMed=8566965;
 RA Schroock E., Badger P., Larson D., Erdos M., Wynshaw-Boris A.,
 RA Ried T., Brody L.;
 RT "The murine homologue of the human breast and ovarian cancer
 RT susceptibility gene Brca1 maps to mouse chromosome 11D.";
 RL Hum. Genet. 97:256-259(1996).
 CC -1- FUNCTION: Plays a central role in DNA repair by facilitating
 CC cellular response to DNA repair. Required for appropriate cell
 CC cycle arrests after ionizing irradiation in both the S-phase and
 CC the G2 phase of the cell cycle. Involved in transcriptional
 CC regulation of p21 in response to DNA damage. May function as a
 CC transcriptional regulator. Mediates E2-dependent ubiquitination (By
 CC similarity).
 CC -1- SUBUNIT: Part of the BRCA1-associated genome surveillance complex
 CC (BASC), which contains BRCA1, MSH2, MSH6, MLH1, ATM, BLM, FMS2 and
 CC the RAD50-MRE11-NBS1 protein complex. This association could be a
 CC dynamic process changing throughout the cell cycle and within
 CC subnuclear domains. CtIP interacts specifically with the BRCT
 CC domains. Interacts with RNA polymerase II holoenzyme. Interacts
 CC with SWC1L1 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- TISSUE SPECIFICITY: In the embryo, expressed in otic vesicles at
 CC day 9.5. At day 10.5, this expression decreases and high levels
 CC are found in the neuroectoderm. At days 11-12.5, high levels in
 CC differentiating keratinocytes and whisker pad primordia. At days
 CC 14-17, expression also observed in kidney epithelial cells. In
 CC the adult, highest levels found in spleen, thymus, lymph nodes,
 CC epithelial organs, and alveolar and ductal epithelial cells of
 CC the mammary gland. Very low levels in brain, kidney, and skin. No
 CC expression in heart, liver or lung.
 CC -1- DEVELOPMENTAL STAGE: In the mammary gland, expression increases
 CC dramatically during pregnancy. Levels fall during lactation and
 CC increase again during post-lactational regression of the
 CC mammary gland.
 CC -1- PTM: Phosphorylated by ATM upon ionizing radiation (By
 CC similarity).
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -1- SIMILARITY: Contains 2 BRCT domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL; U35641; AAB17113.1; -;
 CC EMBL; U31625; AAB17114.1; -;
 CC EMBL; U32446; AAA96393.1; -;
 CC EMBL; U36475; AAC52323.1; -;
 CC EMBL; U31835; AAA97442.1; -;
 CC MGI; MGI:104537; Brca1.
 CC GO; GO:0005737; C:cytoplasm; IDA.
 CC GO; GO:0008274; C:gamma-tubulin ring complex; ISS.
 CC GO; GO:0005634; C:nucleus; ISS.
 CC GO; GO:0016329; F:apoptosis regulator activity; ISS.
 CC GO; GO:0005515; F:protein binding activity; ISS.
 CC GO; GO:0016563; F:transcriptional activator activity; ISS.
 CC GO; GO:0015631; F:tubulin binding activity; ISS.
 CC GO; GO:0006915; P:apoptosis; ISS.
 CC GO; GO:0007098; P:cell cycle; ISS.
 CC GO; GO:0006787; P:DNA damage response, induction of cyclin-de. . .; ISS.
 CC GO; GO:0045739; P:positive regulation of DNA repair; ISS.
 CC GO; GO:0042127; P:regulation of cell proliferation; ISS.
 CC GO; GO:0006559; P:regulation of transcription from Pol III pr. . .; ISS.
 CC Interpro: IPR001357; BRCT.
 CC Interpro: IPR002378; Brst cancer1.
 CC Interpro: IPR001841; Znf_Ting.
 CC Pfam; PF00533; BRCT; 2.
 CC Pfam; PF00097; ZF-C3HC4; 1.
 CC PRINTS; PR00493; BRSTCANCER1.
 CC SMART; SM00292; BRCT; 2.
 CC SMART; SM00184; RING; 1.
 CC PROSITE; PSS0172; BRCT; 2.
 CC PROSITE; PSS00518; ZF_RING_1; 1.
 CC PROSITE; PSS0089; ZF_RING_2; 1.
 CC DNA repair; Zinc-finger; DNA-binding; Nuclear protein; Polymorphism;
 CC Anti-oncogene; Repeat; Phosphorylation.
 CC ZN_FING 24
 CC FT 1585 1679
 CC FT DOMAIN 1698 1797
 CC FT DOMAIN 1562 1567
 CC FT MOD_RES 1343 1343
 CC FT MOD_RES 1481 1481
 CC FT VARIANT 93 93
 CC FT VARIANT 305 305
 CC FT VARIANT 319 319
 CC FT VARIANT 377 377
 CC FT VARIANT 550 550
 CC FT VARIANT 652 652
 CC FT VARIANT 765 765
 CC FT VARIANT 917 917
 CC FT VARIANT 933 933
 CC FT VARIANT 1122 1122
 CC FT VARIANT 1206 1206
 CC FT VARIANT 1212 1212
 CC FT VARIANT 1255 1255
 CC FT VARIANT 1261 1261
 CC FT VARIANT 1264 1264
 CC FT VARIANT 1269 1269
 CC FT VARIANT 1283 1283
 CC FT VARIANT 1337 1337
 CC FT VARIANT 1349 1349
 CC FT VARIANT 1352 1353
 CC FT VARIANT 1381 1381
 CC FT VARIANT 1390 1390
 CC FT VARIANT 1400 1400
 CC P -> L (IN STRAIN 129/SVJ).
 CC T -> S (IN STRAIN 129/SVJ).
 CC Q -> A (IN STRAIN 129/SVJ).
 CC Q -> E (IN STRAIN 129/SVJ).
 CC K -> Q (IN STRAIN 129/SVJ).
 CC K -> P (IN STRAIN 129/SVJ).
 CC S -> P (IN STRAIN 129/SVJ).
 CC P -> L (IN STRAIN 129/SVJ).
 CC P -> L (IN STRAIN 129/SVJ).
 CC C -> S (IN STRAIN 129/SVJ).
 CC K -> I (IN STRAIN 129/SVJ).
 CC S -> R (IN STRAIN 129/SVJ).
 CC RM -> GI (IN STRAIN 129/SVJ).
 CC S -> R (IN STRAIN 129/SVJ).
 CC H -> N (IN STRAIN 129/SVJ).
 CC V -> A (IN STRAIN 129/SVJ).
 CC P -> A (IN STRAIN 129/SVJ).
 CC T -> K (IN STRAIN 129/SVJ).
 CC T -> N (IN STRAIN 129/SVJ).
 CC P -> T (IN STRAIN 129/SVJ).
 CC EG -> OR (IN STRAIN 129/SVJ).
 CC S -> P (IN STRAIN 129/SVJ).
 CC G -> A (IN STRAIN 129/SVJ).
 CC V -> D (IN STRAIN 129/SVJ).

FT VARIANT 1503 1503 E -> Q (IN STRAIN 129/SVJ)
 FT VARIANT 1549 1549 V -> A (IN STRAIN 129/SVJ)
 FT VARIANT 1680 1680 T -> K (IN STRAIN 129/SVJ)
 FT VARIANT 1712 1712 D -> E (IN STRAIN 129/SVJ)
 FT VARIANT 1721 1721 D -> B (IN STRAIN 129/SVJ)
 SQ SEQUENCE 1812 AA; 198669 MW; 2291EA74150BB86A CRC64;

Query Match 8.8%; Score 78; DB 1; Length 1812;
 Best Local Similarity 27.4%; Pred No. 66;
 Matches 32; Conservative 22; Mismatches 37; Indels 26; Gaps 6;

QY 5 IFAFGVHLVSLCSGKA---ICXNGISKRT-----FEIKKEI-----ASCGDVAKAI 50
 DB 42 IFCRCMLKLNQKQKSPCLCKNEITRSLQSTRSQALAEELRLMAAFELDTGMQL 101
 QY 51 INLVYGAQRNSYREL---ALLVTVG-----PRIS-GSKLKAQIMYQNL 95
 DB 102 TNGFSFSKRRNNSCERLNEASIIQSGYRNRRLPOVEPGNATLKDSLGVLSNL 158

RESULT 9

VCOM_ADE05 STANDARD; PRT; 368 AA.
 AC P24938;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Minor core protein (Protein V).
 GN PV.
 OS Human adenovirus type 5.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=28285;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92087470; PubMed=1272603;
 RA Chroboczek J., Bleher F., Jacrot B.;
 RT "The sequence of the genome of adenovirus type 5 and its comparison
 RL with the genome of adenovirus type 2.";
 RL Virology 186:280-285(1992).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M73260; AAA96409.1; -;
 DR PIR: C39449; FOADM5.
 DR InterPro: IPR005608; Adeno_PV.
 DR Pfam: PF03910; Adeno_PV.1.
 KW Core protein; Late protein.
 SQ SEQUENCE 368 AA; 41446 MW; 722E6C6D22C692A4 CRC64;

Query Match 8.8%; Score 77.5; DB 1; Length 368;
 Best Local Similarity 25.4%; Pred No. 11;
 Matches 43; Conservative 21; Mismatches 50; Indels 55; Gaps 8;

QY 29 SKRTPEELKEBIASCGDV---AKAIIINLVYGAQRNSYERIALALVTVGPRLS----- 79
 DB 89 STRTYKRYVDEYGDDELLEQANERLGFPAVK---RKMDMALPLDNGNTPPSLKPVTL 145
 QY 80 -----GSKLKAQIMYQNL-OQGLKVLHLPVRIPIHMERGES 119
 DB 146 QOVLPALAPSEKRGKLGKESGDLAPVQLMWPVKRQLRDVLEKMTVEP-----GHLP 197
 QY 120 AVMLEPRHKLIALILGLS-----SSIGPRPGIATVILVYTS 155
 DB 198 EVRVAP-IKOVA-PGLGVQTVDVQVLPPTSSISIAIATGEMTQSPVAS 244

RESULT 10

AB10_HUMAN STANDARD; PRT; 738 AA.
 ID AB10_HUMAN
 AC Q9NRK6; Q13040; Q9H3V0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ATP-binding cassette, sub-family B, member 10, mitochondrial precursor
 DE (ATP-binding cassette transporter 10) (ABC transporter 10 protein)
 DE (Mitochondrial ATP-binding cassette 2) (M-ABCC2).
 GN ABCB10.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymphoblast;
 RX MEDLINE=20382730; PubMed=10922475;
 RA Zhang F., Hogue D.L., Liu L., Fisher C.L., Hui D., Childs S.,
 RA Ling V.;
 RT "M-ABCC2, a new human mitochondrial ATP-binding cassette membrane
 RT protein.";
 RL FEBS Lett. 478:89-94(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ito K., Suzuki H., Sugiyama Y.;
 RT Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 510-738 FROM N.A.
 RX MEDLINE=95284468; PubMed=7766993;
 RA Allkhets R., Gerard B., Glavac D., Ravnik-Glavac M., Jenkins N.A.,
 RA Gilbert D.J., Copeland N.G., Modi W., Dean M.;
 RT "Characterization and mapping of three new mammalian ATP-binding
 RT transporter genes from an EST database.";
 RL Mamm. Genome 6:114-117(1995).
 RN [4]
 RP VARIANT SER-150.
 RX MEDLINE=21686803; PubMed=11829140;
 RA Saito S., Iida A., Sekine A., Mura Y., Ogawa C., Kawachi S.,
 RA Higuchi S., Nakamura Y.;
 RT "Three hundred twenty-six genetic variations in genes encoding nine
 RT members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the
 RT Japanese population.";
 RL J. Hum. Genet. 47:38-50(2002).
 CC -----
 CC -1- FUNCTION: May mediate critical mitochondrial transport functions
 CC related to heme biosynthesis (By similarity).
 CC -1- SUBUNIT: Homodimer or heterodimer (Potential).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in bone marrow,
 CC expressed at intermediate to high levels in skeletal muscle, small
 CC intestine, thyroid, heart, brain, placenta, liver, pancreas,
 CC prostate, testis, ovary, leukocyte, stomach, spinal cord, lymph
 CC node, trachea and adrenal gland, and low levels are found in lung,
 CC kidney, spleen, thymus and colon.
 CC -----
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF216833; AAF78198.1; -;
 DR EMBL: AB013380; BAB20265.1; -;
 DR EMBL: U18237; AA84438.1; -;
 DR Genew; HGNC:41; ABCB10.
 DR MTM; 605454; -;
 DR GO; GO:0005739; C:mitochondrion; NAS.
 DR GO; GO:0005524; F:ATP binding activity; NAS.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; NAS.

Query	Subject	Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
GO:0006810; P:transport; NAS	GO:0006810; P:transport; NAS	100%	76.5	DB 1	738	44	31	82	43	8
InterPro: IPR003593; AAA ATPase	InterPro: IPR003593; AAA ATPase	100%	76.5	DB 1	738	44	31	82	43	8
InterPro: IPR001140; ABC_TM_transp	InterPro: IPR001140; ABC_TM_transp	100%	76.5	DB 1	738	44	31	82	43	8
InterPro: IPR003439; ABC_transporter	InterPro: IPR003439; ABC_transporter	100%	76.5	DB 1	738	44	31	82	43	8
Pfam: PF00664; ABC_membrane; 1	Pfam: PF00664; ABC_membrane; 1	100%	76.5	DB 1	738	44	31	82	43	8
Pfam: PF00005; ABC_tran; 1	Pfam: PF00005; ABC_tran; 1	100%	76.5	DB 1	738	44	31	82	43	8
ProDom: PD00006; ABC_transporter; 1	ProDom: PD00006; ABC_transporter; 1	100%	76.5	DB 1	738	44	31	82	43	8
SMART: SM00382; AAA_1	SMART: SM00382; AAA_1	100%	76.5	DB 1	738	44	31	82	43	8
PROSITE: PS00211; ABC_TRANSPORTER_1; 1	PROSITE: PS00211; ABC_TRANSPORTER_1; 1	100%	76.5	DB 1	738	44	31	82	43	8
PROSITE: PS50893; ABC_TRANSPORTER_2; 1	PROSITE: PS50893; ABC_TRANSPORTER_2; 1	100%	76.5	DB 1	738	44	31	82	43	8
KW ATP-binding; Transmembrane; Transport; Mitochondrion; Inner membrane; Transist peptide; Polymorphism	KW ATP-binding; Transmembrane; Transport; Mitochondrion; Inner membrane; Transist peptide; Polymorphism	100%	76.5	DB 1	738	44	31	82	43	8
TRANSIT	TRANSIT	100%	76.5	DB 1	738	44	31	82	43	8
CHAIN	CHAIN	100%	76.5	DB 1	738	44	31	82	43	8
DOMAIN	DOMAIN	100%	76.5	DB 1	738	44	31	82	43	8
TRANSMEM	TRANSMEM	100%	76.5	DB 1	738	44	31	82	43	8
TRANSMEM	TRANSMEM	100%	76.5	DB 1	738	44	31	82	43	8
TRANSMEM	TRANSMEM	100%	76.5	DB 1	738	44	31	82	43	8
DOMAIN	DOMAIN	100%	76.5	DB 1	738	44	31	82	43	8
TRANSMEM	TRANSMEM	100%	76.5	DB 1	738	44	31	82	43	8
TRANSMEM	TRANSMEM	100%	76.5	DB 1	738	44	31	82	43	8
DOMAIN	DOMAIN	100%	76.5	DB 1	738	44	31	82	43	8
NP_BIND	NP_BIND	100%	76.5	DB 1	738	44	31	82	43	8
VARIANT	VARIANT	100%	76.5	DB 1	738	44	31	82	43	8
CONFLICT	CONFLICT	100%	76.5	DB 1	738	44	31	82	43	8
CONFLICT	CONFLICT	100%	76.5	DB 1	738	44	31	82	43	8
CONFLICT	CONFLICT	100%	76.5	DB 1	738	44	31	82	43	8
CONFLICT	CONFLICT	100%	76.5	DB 1	738	44	31	82	43	8
CONFLICT	CONFLICT	100%	76.5	DB 1	738	44	31	82	43	8
SEQUENCE	SEQUENCE	100%	76.5	DB 1	738	44	31	82	43	8
Query Match	Query Match	100%	76.5	DB 1	738	44	31	82	43	8
Best Local Similarity	Best Local Similarity	100%	76.5	DB 1	738	44	31	82	43	8
Matches	Matches	100%	76.5	DB 1	738	44	31	82	43	8
Conservative	Conservative	100%	76.5	DB 1	738	44	31	82	43	8
Mismatches	Mismatches	100%	76.5	DB 1	738	44	31	82	43	8
Indels	Indels	100%	76.5	DB 1	738	44	31	82	43	8
Gaps	Gaps	100%	76.5	DB 1	738	44	31	82	43	8

```

XX NCBI_TaxID=6666;
RN SEQUENCE FROM N.A.
RP STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Hatt D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettein H., Richardson D.,
RA Ermolaeva M.D., Vanathavan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA Mcnoldale L., Uterachavan T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000) .
CC -I- FUNCTION: CATALYZES THE TRANSFER OF A METHYL GROUP FROM 5-
CC METHYLTERADYNOFORATE TO HOMOCYSTEINE RESULTING IN METHIONINE
CC FORMATION (BY SIMILARITY) .
CC -I- CATALYTIC ACTIVITY: 5-methyltetrahydropteroyltri-L-glutamate + L-
CC homocysteine = tetrahydropteroyltri-L-glutamate + L-methionine.
CC -I- COFACTOR: ZINC, BINDS ONE ION PER SUBUNIT (BY SIMILARITY) .
CC -I- PATHWAY: TERMINAL STEP IN THE DE NOVO BIOSYNTHESIS OF METHIONINE.
CC -I- SIMILARITY: Belongs to the vitamin-B12 independent methionine
CC synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC The European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch) .
CC -----
CC EMBL; AE004248; AAF94854.1; -.
CC DR PIR; E82167; E82167.
CC DR TIGR; VC1704; -.
CC DR HAMAP; MF_00172; -.
CC DR InterPro; IPR006276; Met_syn_B12ind.
CC DR InterPro; IPR002629; Methionine_synth.
CC DR Pfam; PF01717; Methionine_synth; 1.
CC DR Prodom; PD004692; Methionine_synth; 2.
CC DR TIGRFAMs; TIGR01371; met_syn_B12ind; 1.
CC KM Transferrase; Methyltransferase; Methionine biosynthesis; Zinc; Repeat;
CC Complete proteome.
CC KW METAL 646 646 ZINC (BY SIMILARITY) .
CC FT METAL 648 648 ZINC (BY SIMILARITY) .
CC FT METAL 731 731 ZINC (BY SIMILARITY) .
CC FT SEQUENCE 761 AA; 85046 MW; C65094CF7989F00C CRC64;
SQ
Query Match 8.7%; Score 76.5; DB 1; Length 761;
Best Local Similarity 25.4%; Pred. No. 32;
Matches 34; Conservative 21; Mismatches 42; Indels 37; Gaps 7;
OY 33 FEETKEEELASGCVDAKATINLAVY---GKAQNRBYERALLVDYVGRPLSSKMLEKIQ 89
DB 145 FEEVVEALQADHDVDPVLLGPISYLVLGKEVEEGRDRLTL-----PRLT-----TAYQ 193
OY 90 IMYQNLQODDGLGEKYLH-EPVYIPRMHGESESAWMLBP---IHKIALIGLSSSIGTPREG 145
DB 194 AILSKLAQGVQWQIDBPI-----LATELEPRWQEPFLAYQVINGDV----- 237
OY 146 ITAEVLVTSFDEL 159
DB 238 ---KLLLTTFDSV 248

```


DE DNA replication licensing factor CDC47 (Cell division control protein
 47).
 GN CDC47 OR YBR202W OR YBR1441.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CC NCBI_TaxID=4932;
 CC [1]
 RN SEQUENCE FROM N.A.
 RA Dalton S.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=93377417; PubMed=8368014;
 RA Busserreau F., Maillet L., Gallion L., Jacquet M.;
 RT "A 12.8 kb segment, on the right arm of chromosome II from
 RT Saccharomyces cerevisiae including part of the DUR1,2 gene, contains
 RT five putative new genes.";
 RL Yeast 9:797-806(1993).
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U14730; AAA6309.1; -;
 DR EMBL; Z21487; CA79689.1; -;
 DR EMBL; Z35071; CA85166.1; -;
 DR PIR; S34027; S34027.
 DR GK; P38132; -;
 DR SCD; S0000406; CDC47.
 DR GO; GO:0005737; C:nucleus; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005656; C:pre-replicative complex; IDA.
 DR GO; GO:0006267; P:pre-replicative complex formation and maint. . .; IPI.
 DR InterPro; IPR003593; AAA_Atpase.
 DR InterPro; IPR001208; MCM.
 DR Pfam; PF00493; MCM; 1.
 DR ProDom; PD001041; MCM; 1.
 DR SMART; SM00382; AAA; 1.
 DR SMART; SM00350; MCM; 1.
 DR PROSITE; PS00847; MCM_1; 1.
 DR PROSITE; PS50051; MCM_2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 KM DNA replication; Cell cycle; ATP-binding.
 FT DOMAIN 410 617 MCM.
 FT NP_BIND 460 467 ATP (POTENTIAL).
 FT CONFLICT 552 552 G -> V (IN REF. 2).
 FT CONFLICT 556 558 TLN -> NPG (IN REF. 2).
 FT CONFLICT 574 574 Y -> I (IN REF. 2).
 SQ SEQUENCE 845 AA; 94942 MW; ADA66C719D96B4A CRC64;

Query Match. 8.7%; Score 76.5; DB 1; Length 845;
 Best Local Similarity 23.1%; Pred. No. 36;
 Matches 40; Conservative 34; Mismatches 56; Indels 43; Gaps 8;

QY 11 GVHLLSLGSKAICNGISKRTFEI-----KERIAS---CGDVAKAIINLNAVYGRKON 61
 DB 353 GILPAPYTGFKALAGLTETYLEAQPVRQHKKFFAFSLTSDVEERVMELITSGDVYN 412
 QY 62 RSYERIALVDTVGPRLSGNLEKAIQIMYONLQODGLEKVLHPVRIPIHMERGESAV 121
 DB 413 R-----LAKSIAPETIGNDVKKALL---LIVGVDKRVGGMKI-----RGDINVC 457
 QY 122 ML-EPRIHKIALI-----GLSSIGTPPGITAEVLVTSFDEL 159
 DB 458 LMGDPGVAKSGQLTKAIKISPRGVYTTGSS-----GVLGTFAAVKKDPVTDEM 506

RESULT 13
 Y694_METUA STANDARD; PRT; 414 AA.
 ID Y694_METUA
 AC Q58105;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein M0694.
 GN M0694.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CC Methanocaldococcaceae; Methanocaldococcus.
 CC NCBI_TaxID=2190;
 CC [1]
 RN SEQUENCE FROM N.A.
 RA STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kierlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Utermack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: BELONGS TO THE NOP5/NOP56 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U67516; AAB98689.1; -;
 DR PIR; F64386; F64386.
 DR TIGR; M0694; -;
 DR InterPro; IPR002687; NOP.
 DR Pfam; PF01798; NOP; 1.
 DR ProDom; PD004104; NOP; 1.
 KW Hypothetical protein; Complete proteome.
 KM Hypothetical protein; ASP/GLU/LYS-RICH.
 FT DOMAIN 349 414
 SQ SEQUENCE 414 AA; 47799 MW; A9092EFC3C82C407 CRC64;

Query Match. 8.6%; Score 75.5; DB 1; Length 414;
 Best Local Similarity 23.3%; Pred. No. 19;
 Matches 47; Conservative 29; Mismatches 59; Indels 67; Gaps 11;

QY 23 ICKNGISKRTFEIERKERIAS---CGDVAKAIIN-----LAVYGAQNRSYER 66
 DB 172 ITKGKRNKTKSGQLKILSKLAGKTAFAAKSMGSELDYDLDVYVKAPEINHLIYK 231
 QY 67 LALLVD-----TVGPRISG-SKNLEKAIQIMYONLQODGLEK---V 103
 DB 232 RKEIYVYLEKLMNEADNITKLAGVSLGARLIGLGLGELKIALPASTIVLGAEKALFA 291
 QY 104 HL-----EPVR-----IPHMERGESAVMLEPRIHKIALILGLSSIGTPPGIT 147
 DB 292 HLRNGVPPRGIIYVHPLIQGSPHWRGKIARL-----ACKLAIARADYVG---DYIA 344
 QY 148 AEVLVTSFDELORASEARGK 169
 DB 345 DELL-----EKLNRVVEIRRK 361

RESULT 14
 YK04_YEAST

ID YK04_YEAST STANDARD; PRT; 579 AA.
AC P32343;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 65.1 kDa protein in RKN3-SRP21 intergenic region.
GN YK124W OR YK1529.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-120 FROM N.A.
RA Rad M.R., Xu G., Kirchath L., Fritz C., Keuchel H., Hollenberg C.P.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 41-579 FROM N.A.
RX MEDLINE=9283952; PubMed=1514329;
RA Colledge L., Richard G.-F., Thierry A., Dujon B.;
RT "Sequence of a segment of Yeast chromosome XI identifies a new
RT mitochondrial carrier, a new member of the G protein family, and a
RT protein with the PAKK motif of the H1 histones."
RL Yeast 8:325-336(1992).
CC -1- SIMILARITY: TO YEAST YMR171C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; Z28123; CA81964.1; -
DR EMBL; S44213; AAB23074.1; -
DR PIR; S37953; S37953.
DR SGD; S0001607; SSH4.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00623; SPRY; 1.
DR SMART; SM00449; SPRY; 1.
KW Hypothetical protein.
SQ SEQUENCE 579 AA; 65060 MM; 700FICBA0749754 CRC64;
Query March 8.6%; Score 75.5; DB 1; Length 579;
Best Local Similarity 24.1%; Pred. No. 28;
Matches 32; Conservative 24; Mismatches 64; Indels 13; Gaps 4;
QY 45 DVAKAI-----INLAVYGKQNRSEYRLALVD---TWGPRLSGSKNLEKAIQIMYQNTL 95
DB 321 DVTQNIIGDIFIGAFNAATRTYTRGDLLEDPNVSFRALBEGKDIEVAKDLQRVHD 380
QY 96 QQDGLKXVLEPRVIRPWEGESAVMLKPRIKIAIIGLSSIGTPPEGITAEV---LV 152
DB 381 PHDESDEMTSEBEL-HVNLGQVGFVEANVKYAFSGSVYGOIGIPAYNGTEIKKQTI 439
QY 153 VTSFDELQRRASE 165
DB 440 LQKGEELPRYAD 452
RESULT 15
SMC4_MOUSE
ID SMC4_MOUSE STANDARD; PRT; 1286 AA.
AC Q8CG47; Q8BT57; Q8BT59; Q99K21;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Structural maintenance of chromosomes 4-like 1 protein (Chromosome-
DE associated polypeptide C) (XCAP-C homolog).
GN SMC4L1 OR SMC4 OR CAPC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Cobbe N., Heck M.M.S.;
RT "Phylogenetic analysis of SMC proteins."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-602 FROM N.A.
RC SRRATN=C57BL/6J; and NOD; TISSUE=Embryonic liver, and Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nakaido I., Osato N., Saito R., Suzuki H., Yamada I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schirral L.M., Kanapin A., Matsuda H., Batalov S., Batzel K.W.,
RA Blake J.A., Brad D., Brasic V., Chochua C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Fraser K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Saitana A., Schneider C., Semple C.A., Setton M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Velardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yamagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hitzonen-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shimagawa A.,
RA Yaunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE OF 1103-1286 FROM N.A.
RC TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stedman M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Bawa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shcherbenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16993(2002).
CC -1- FUNCTION: Central component of the condensin complex, a complex
CC required for conversion of interphase chromatin into mitotic-like
CC condensed chromosomes. The condensin complex probably introduces
CC positive supercoils into relaxed DNA in the presence of type I
CC topoisomerases and converts nicked DNA into positive knotted forms
CC in the presence of type II topoisomerases (By similarity).
CC -1- SUBUNIT: Forms a heterodimer with SMC2L1. Component of the
CC condensin complex, which contains the SMC2L1 and SMC4L1
CC heterodimer, and three non SMC subunits that probably regulate the
CC complex: BRN1/CAPB, CNA1/CAP2 and CAPG (By similarity).

CC -|- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. In interphase
CC cells, the majority of the condensin complex is found in the
CC cytoplasm, while a minority of the complex is associated with
CC chromatin. A subpopulation of the complex however remains
CC associated with chromosome foci in interphase cells. During
CC mitosis, most of the condensin complex is associated with the
CC chromatin. At the onset of prophase, the regulatory subunits of
CC the complex are phosphorylated by CDC2, leading to condensin's
CC association with chromosome arms and to chromosome condensation.
CC Dissociation from chromosomes is observed in late telophase (By
CC similarity).
CC -|- DOMAIN: The hinge domain, which separates the large intramolecular
CC coiled coil regions, allows the heterodimerization with SMC2L1,
CC forming a V-shaped heterodimer (By similarity).
CC -|- SIMILARITY: Belongs to the SMC family. SMC4 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AJ534940; CAD59183.1; -
CC EMBL; AK088846; BAC40608.1; -
CC EMBL; AK088350; BAC40297.1; -
CC EMBL; BC005507; AA05507.1; -
CC MGI; 1917349; Smc411.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR003405; SMC_C.
CC InterPro; IPR003395; SMC_N.
CC Pfam; PF02483; SMC_C; 1.
CC Pfam; PF02463; SMC_N; 1.
CC DNA condensation; Mitosis; Cell cycle; ATP-binding; Coiled coil;
KW Nuclear protein.
KW NP_BIND 111 118 ATP (POTENTIAL).
FT DOMAIN 270 589 COILED COIL (POTENTIAL).
FT DOMAIN 590 767 FLEXIBLE HINGE.
FT DOMAIN 768 1018 COILED COIL (POTENTIAL).
FT DOMAIN 1068 1133 COILED COIL (POTENTIAL).
FT DOMAIN 1189 1224 ALA/ASP-RICH (DA-BOX).
FT DOMAIN 64 69 POLY-PRO.
SQ SEQUENCE 1286 AA; 146894 MW; CFEAD84199C3CEB5 CRC64;

Query Match 8.6%; Score 75.5; DB 1; Length 1286;
Best local Similarity 22.8%; Pred. No. 74;
Matches 41; Conservative 37; Mismatches 61; Indels 41; Gaps 12;

QY 16 SLCSGKAICKNGISRTFEIEKEIASCGDAKAIINLAVYGKAKNRSYERIALLVDTVG 75
Db 938 SVCRTEKIKD--TEKEINDKTEIAKNIEDKAEVIN-----NTKTAETSL----- 981

QY 76 PRLSGS-KULEKALQIMYN---LQDGLF-KVHLEPV-----RIPMERGESAV 121
Db 982 PEIQKEHNNLQELKVIQENHALQKDALSIKLEQIDGHISENSKITKWOK-EISKI 1040

QY 122 MLEP---RIKIALIGLSSIGTP-PEGITAEVLVTSFDELQRRASEAR-GKIVVYNQ 175
Db 1041 KLHPVEDNPVETVAVLQSEELERAIKNPSTINETIAL---EAQCREKMPNIGALAEYKK 1096

Search completed: December 22, 2003, 11:50:02
Job time : 13.0896 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:37:00 ; Search time 26.507 Seconds
(without alignments)
1703.674 Million cell updates/sec

Title: US-09-745-763-36_COPY_1_175
Perfect score: 883
Sequence: 1 MKFLIFAFGCVHLSLCSG.....FDLQRRASRARKIVVYQ 175

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	883	100.0	472.4	Q9Y646	Q9Y646 homo sapien
2	883	100.0	541.4	Q9Y5X6	Q9Y5X6 homo sapien
3	872	98.8	472.4	Q8NB21	Q8NB21 homo sapien
4	736	83.4	433.11	Q70216	Q70216 mus musculu
5	736	83.4	470.11	Q9WVJ3	Q9WVJ3 mus musculu
6	734	83.1	472.11	Q9JUV0	Q9JUV0 ratius norv
7	727	82.3	472.11	Q9Z1Y1	Q9Z1Y1 ratius norv
8	493	55.8	140.4	Q9UNM8	Q9UNM8 homo sapien
9	217	24.6	493.5	Q76552	Q76552 acanthocheil
10	180.5	20.4	414.16	Q8B8H7	Q8B8H7 shewanella
11	163.5	18.5	467.16	Q9A5A8	Q9A5A8 caulobacter
12	121	13.7	472.16	Q8P8E8	Q8P8E8 xanthomonas
13	117	13.3	375.16	Q8P6Z5	Q8P6Z5 xanthomonas
14	93.5	10.6	557.16	Q9WZ59	Q9WZ59 thermotoga
15	92.5	10.5	1295.5	Q22257	Q22257 caenorhabdi
16	92	10.4	920.16	Q9PL78	Q9PL78 chlamydia m

17	91.5	10.4	731.2	Q8RT29	Q8RT29 pseudoalter
18	89.5	10.1	215.16	Q9KJ34	Q9KJ34 bacillus ha
19	89	10.1	592.6	Q9XSR3	Q9XSR3 canis fami1
20	87	9.9	449.2	Q93EJ5	Q93EJ5 bacillus 11
21	86.5	9.8	770.17	Q8YPI1	Q8YPI1 mechanopyru
22	85.5	9.7	388.16	Q8ZHQ2	Q8ZHQ2 yersinia pe
23	85.5	9.7	729.2	Q9KH34	Q9KH34 antarctic b
24	85	9.6	349.2	Q8VCG4	Q8VCG4 spiroplasma
25	84.5	9.6	326.16	Q9XAG1	Q9XAG1 streptomyce
26	84.5	9.6	551.12	Q66601	Q66601 epizootic h
27	84.5	9.6	765.16	P72735	P72735 synchocyst
28	82.5	9.3	350.16	Q8X8V6	Q8X8V6 escherichia
29	82	9.3	524.4	Q9NKG3	Q9NKG3 homo sapien
30	82	9.3	593.16	Q8N152	Q8N152 homo sapien
31	82	9.3	821.16	Q55662	Q55662 synchocyst
32	82	9.3	864.10	Q9LZ19	Q9LZ19 arabidopsis
33	81.5	9.2	162.16	Q25207	Q25207 helicobacte
34	81.5	9.2	423.6	Q28014	Q28014 bradyphus tr
35	81.5	9.2	456.2	Q8KJH9	Q8KJH9 rhizobium e
36	81.5	9.2	551.12	Q66600	Q66600 epizootic h
37	81	9.2	579.10	Q941N0	Q941N0 hordium vul
38	81	9.2	1170.5	Q22624	Q22624 caenorhabdi
39	80	9.1	1479.2	Q937H2	Q937H2 klebsiella
40	80	9.1	1486.16	Q8XDG0	Q8XDG0 escherichia
41	80	9.1	1486.16	Q8FJA2	Q8FJA2 escherichia
42	79.5	9.0	377.16	Q8COD7	Q8COD7 staphylococ
43	79.5	9.0	409.16	Q95KU5	Q95KU5 cholesopus d
44	79	8.9	428.17	Q58181	Q58181 pyrococcus
45	79	8.9	499.17	Q26204	Q26204 methanobact

ALIGNMENTS

RESULT 1

Q9Y646 PRELIMINARY; PRT; 472 AA.
ID Q9Y646
AC Q9Y646;
DC 01-NOV-1999 (TREMURel. 12, Created)
DT 01-NOV-1999 (TREMURel. 12, Last sequence update)
DT 01-MAR-2003 (TREMURel. 23, Last annotation update)
DE Aminopeptidase (Plasma glutamate carboxypeptidase).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Liu C.H., Lin B.Y., Chang L.Y.;
RT "Cloning of the human aminopeptidase gene";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF107834; AAD43214.1; -;
DR EMBL; BC020689; AAR20689.1; -;
DR MEROPS; M28.014; -;
DR InterPro; IPR00137; PA.
DR PROSITE; PS00840; PA; 1.
KW Carboxypeptidase.
SQ SEQUENCE 472 AA; 51887 MW; EB6CBD2149E042BF CRC64;

Query Match 100.0%; Score 883; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.9e-75;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKFLIFAFGCVHLSLCSGKAICKNGISKRTREIKKEELASGCDVAKAIINLAVYKQ 60
Db 1 MKFLIFAFGCVHLSLCSGKAICKNGISKRTREIKKEELASGCDVAKAIINLAVYKQ 60

```

Qy 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Db 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Qy 121 VMLEPRHKAIIILGLSSIGTPPEGITAEVLVVTSPFDELORRASEARNGKIIVYVNO 175
Db 121 VMLEPRHKAIIILGLSSIGTPPEGITAEVLVVTSPFDELORRASEARNGKIIVYVNO 175

RESULT 2
Qy5X6 PRELIMINARY; PRT; 541 AA.
ID Qy5X6
AC Qy5X6;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DB Blood plasma glutamate carboxypeptidase precursor (EC 3.4.17.21).
GN PGCP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=99223495; PubMed=10206990;
RX Gingras R., Richard C., El-Alfy M., Morales C.R., Potier M.,
RA Pehezhetsky A.V.;
RT "Purification, cDNA cloning, and expression of a new human blood
RT plasma glutamate carboxypeptidase homologous to N-acetyl-aspartyl-
RT alpha-glutamate carboxypeptidase/prostate-specific membrane antigen.";
RL J. Biol. Chem. 274:11742-11750(1999).
DR EMBL; AF119386; AAD31418.1; -.
DR MEROPS; M28.014; -.
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
KW Carboxypeptidase; Hydrolase; Signal.
FT SIGNAL 1 24
FT CHAIN 45 541 BLOOD PLASMA GLUTAMATE CARBOXYPEPTIDASE.
SQ SEQUENCE 541 AA; 59931 MW; 98138FEE97081F6B CRC64;

Query Match 100.0%; Score 883; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 2.3e-75;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Db 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Qy 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Db 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Qy 121 VMLEPRHKAIIILGLSSIGTPPEGITAEVLVVTSPFDELORRASEARNGKIIVYVNO 175
Db 121 VMLEPRHKAIIILGLSSIGTPPEGITAEVLVVTSPFDELORRASEARNGKIIVYVNO 175

RESULT 3
Q8NBZ1 PRELIMINARY; PRT; 472 AA.
ID Q8NBZ1
AC Q8NBZ1;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DB Hypothetical protein FLJ90651.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RA Isegai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,

```

```

RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto Y., Makamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuho J., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075132; BAC11423.1; -.
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 51784 MW; 83F25966F5DC52P3 CRC64;

Query Match 98.8%; Score 872; DB 4; Length 472;
Best Local Similarity 98.9%; Pred. No. 2.1e-74;
Matches 173; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Db 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Qy 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Db 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Qy 121 VMLEPRHKAIIILGLSSIGTPPEGITAEVLVVTSPFDELORRASEARNGKIIVYVNO 175
Db 121 VMLEPRHKAIIILGLSSIGTPPEGITAEVLVVTSPFDELORRASEARNGKIIVYVNO 175

RESULT 4
O70216 PRELIMINARY; PRT; 433 AA.
ID O70216
AC O70216;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DB Hematopoietic lineage switch 2.
GN Hematopoietic lineage switch 2.
GC PGCP OR HUS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Williams J.H., Chan C.-Y., Klinken S.P.;
RT "Hematopoietic lineage switch 2 (HUS2), a novel mRNA species induced
RT during an erythroid to myeloid lineage switch.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF009513; AAC17945.1; -.
DR MEROPS; M28.014; -.
DR MGD; MGI:1889205; PgcP.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR003137; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS50840; PA; 1.
SQ SEQUENCE 433 AA; 47858 MW; 89C9B93712004669 CRC64;

Query Match 83.4%; Score 736; DB 11; Length 433;
Best Local Similarity 85.7%; Pred. No. 1.1e-61;
Matches 150; Conservative 8; Mismatches 15; Indels 2; Gaps 1;

Qy 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Db 1 MRSFLFIFT--VHLLALSGKAVKNGVSOPTREIREEIANVEDVAKAIINLAVYKQY 58
Qy 61 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Db 59 NRSYERLALLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 118
Qy 121 VMLEPRHKAIIILGLSSIGTPPEGITAEVLVVTSPFDELORRASEARNGKIIVYVNO 175
Db 119 VMLEPRHKAIIILGLSSIGTPPEGITAEVLVVTSPFDELORRASEARNGKIIVYVNO 173

```

RESULT 5
Q9WVJ3 PRELIMINARY; PRT; 470 AA.
AC Q9WVJ3;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Aminopeptidase (Similar to plasma glutamate carboxypeptidase).
GN PCCP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Liu C.H., Lin B.Y., Chang L.Y.;
RT "Cloning of the mouse, aminopeptidase gene."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body, and Mesonephros;
RX MEDLINE=22354685; PubMed=12466851;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RL "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AF107835; AAD43215.1; -
DR EMBL: BC037067; AAH37067.1; -
DR EMBL: AK032972; BAC28105.1; -
DR EMBL: AK075686; BAC35891.1; -
DR MEROPS: M28.014; -
DR MGD: MGI:1889205; PgcP.
DR InterPro: IPR001464; Annexin.
DR InterPro: IPR003137; PA.
DR PROSITE: PS00223; ANNEXIN, 1.
DR PROSITE: PS50840; PA; 1.
KW Carboxypeptidase.
SQ SEQUENCE 470 AA; 51813 MW; 0F3490681691866A CRC64;
Query Match 83.4%; Score 736; DB 11; Length 470;
Best Local Similarity 85.7%; Pred. No. 1.8e-61;
Matches 150; Conservative 8; Mismatches 15; Indels 2; Gaps 1;
QY 1 MKFLIFAFPGVHLLSCGKAIKNGISKRTPEIKETIASCGDVAKAIINLAVYGAQ 60
DB 1 MRSIFFLFVAVVHLFSLGSGKAIYKSGVSGRTPEIKETIANVEDVAKAIINLAVYGYQ 58
QY 61 NRSYERLALVDVTVGPRLSGSKNLEKAIQIMYONLQODGLENVHLEQVRIPHWGREGESA 120
DB 59 NRSYERLGLVDVTVGPRLSGSKNLEKAIQIMYONLQODGLENVHLEQVRIPHWGREGESA 118
QY 121 VMEPRIHKIALIGLSSIGTPEGITAEVLVVSFDELQRRASEARAGKIIVNQ 175
DB 119 VMEPRIHKIALIGLSSIGTPEGITAEVLVVSFDELQRRASEARAGKIIVNQ 173
RESULT 6
Q9JLVO PRELIMINARY; PRT; 472 AA.
AC Q9JLVO;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Liver annexin-like protein.
GN LAL.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Servillo G., Della Fazio M.A., Piobbico D., Bartoli D., Castellani M.,
RA Brancorini S., Viola Magni M.,
RT "LAL, a novel gene involved during liver regeneration."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF131077; AAF36518.1; -
DR MEROPS: M28.014; -
DR InterPro: IPR001464; Annexin.
DR InterPro: IPR003137; PA.
DR PROSITE: PS00223; ANNEXIN, 1.
DR PROSITE: PS50840; PA; 1.
SQ SEQUENCE 472 AA; 52011 MW; 55C68FB90E63F265 CRC64;
Query Match 83.1%; Score 734; DB 11; Length 472;
Best Local Similarity 85.1%; Pred. No. 2.8e-61;
Matches 149; Conservative 8; Mismatches 18; Indels 0; Gaps 0;
QY 1 MKFLIFAFPGVHLLSCGKAIKNGISKRTPEIKETIASCGDVAKAIINLAVYGAQ 60
DB 1 MRSIFFLFVAVVHLFSLGSGKAIYKSGVSGRTPEIKETIANVEDVAKAIINLAVYGYQ 60
QY 61 NRSYERLALVDVTVGPRLSGSKNLEKAIQIMYONLQODGLENVHLEQVRIPHWGREGESA 120
DB 61 NRSYERLGLVDVTVGPRLSGSKNLEKAIQIMYONLQODGLENVHLEQVRIPHWGREGESA 120
QY 121 VMEPRIHKIALIGLSSIGTPEGITAEVLVVSFDELQRRASEARAGKIIVNQ 175
DB 121 VMEPRIHKIALIGLSSIGTPEGITAEVLVVSFDELQRRASEARAGKIIVNQ 175
RESULT 7
Q9Z1Y1 PRELIMINARY; PRT; 472 AA.
AC Q9Z1Y1;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hematopoietic lineage switch 2 related protein.
GN HLS2-RP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fisher;
RA Chen Y., Talmage D.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF097723; AAC72384.1; -
DR MEROPS: M28.014; -
DR InterPro: IPR001464; Annexin.
DR InterPro: IPR003137; PA.
DR PROSITE: PS00223; ANNEXIN, 1.
DR PROSITE: PS50840; PA; 1.
SQ SEQUENCE 472 AA; 51969 MW; DC2111651879E6E2 CRC64;
Query Match 82.3%; Score 727; DB 11; Length 472;
Best Local Similarity 84.6%; Pred. No. 1.3e-60;
Matches 148; Conservative 8; Mismatches 19; Indels 0; Gaps 0;
QY 1 MKFLIFAFPGVHLLSCGKAIKNGISKRTPEIKETIASCGDVAKAIINLAVYGAQ 60
DB 1 MRSIFFLFVAVVHLFSLGSGKAIYKSGVSGRTPEIKETIANVEDVAKAIINLAVYGYQ 60
QY 61 NRSYERLALVDVTVGPRLSGSKNLEKAIQIMYONLQODGLENVHLEQVRIPHWGREGESA 120
DB 61 NRSYERLGLVDVTVGPRLSGSKNLEKAIQIMYONLQODGLENVHLEQVRIPHWGREGESA 120

Qy 121 VMLERIRIKHAIILGSSIGTPPEGITAEVLVVTSPFDELQRRASERARKIVVYQ 175
 Db 121 VVVVPRHKKAIILGSSIGTPPEGITAEVLVVTSPFDELQRRASERARKIVVYQ 175

RESULT 8

ID Q9UNM8 PRELIMINARY; PRT; 140 AA.
 AC Q9UNM8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Aminopeptidase (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Liu C.H., Lin B.Y., Chang L.Y.;
 RT "Cloning of the human aminopeptidase gene."
 RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF107833; AAD43213.1;
 DR MEROPS; M28.014;
 DR InterPro; IPR003137; PA.
 DR PROSITE; PSS0840; PA; 1.
 FT NON TER 1
 FT NON TER 140
 SQ SEQUENCE 140 AA; 15472 MW; F34ACEAB33A1AD24 CRC64;

Query Match 55.8%; Score 493; DB 4; Length 140;
 Best Local Similarity 100.0%; Pred. No. 4,6e-39; Indels 0; Gaps 0;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 LSGSKNLEKAIQIMYQNLQODGLEKYLHPVIRPHMERGESAVMLEPRIRIKHAIILGGS 137
 Db 1 LSGSKNLEKAIQIMYQNLQODGLEKYLHPVIRPHMERGESAVMLEPRIRIKHAIILGGS 60
 Qy 138 SIGTPPEGITAEVLVVTSPFDELQRRASERARKIVVYQ 175
 Db 61 SIGTPPEGITAEVLVVTSPFDELQRRASERARKIVVYQ 98

RESULT 9

ID Q76552 PRELIMINARY; PRT; 493 AA.
 AC Q76552;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Aminopeptidase ES-62 precursor.
 OS Acanthocheilichthys vitreum (Dipetalonema vitreum).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Acanthocheilichthys.
 NCBI_TaxID=6277;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20055923; PubMed=10589978;
 RA Harnett W., Houston K.M., Tate E., Garate T., Apfel H., Adam R.,
 RA Haslam S.M., Panico M., Paxton T., Dell A., Morris H., Brzeski H.;
 RT "Molecular cloning and demonstration of an aminopeptidase activity in
 a filarial nematode glycoprotein."
 RL Mol. Biochem. Parasitol. 104:11-23(1999).
 DR EMBL; AF077194; AAC28365.1;
 DR MEROPS; M28.015;
 DR InterPro; IPR003137; PA.
 DR PROSITE; PSS0840; PA; 1.
 KW Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 493 AMINOPEPTIDASE ES-62.
 SQ SEQUENCE 493 AA; 54916 MW; C7E9227DD18E1DFA CRC64;

Query Match 24.6%; Score 217; DB 5; Length 493;
 Best Local Similarity 32.5%; Pred. No. 4,2e-12;
 Matches 55; Conservative 32; Mismatches 58; Indels 24; Gaps 5;

Qy 15 LSLCSKAIC-----KNGISKRTPEIRKEIASCGDVAKAIINLVYKQNRSTYRL 67
 Db 14 LTVVLGAADVLPDKTAPKKNYIOETFGKEVAB-----LIQYTKBEVGLAYQWL 62
 Qy 68 ALIVDTVGPRLSSKNLEKAIQIMYQNLQODGLEKYLHPVIRPHMERGESAVMLEPR 126
 Db 63 SKLVDFGHRNVGSDSLKSIAPLESKLDNDFKVTETVPPNPHVVRGNDVEMTEPR 122
 Qy 127 IHKIAIILGSSIGTPPEGITAEVLVVTSPFDELQRRASERARKIVVYQ 175
 Db 123 NQRLNVLALG---GSEPSATGAVTVYDDDV--KDDVRGKIVVYQ 166

RESULT 10

ID Q8EBH7 PRELIMINARY; PRT; 414 AA.
 AC Q8EBH7;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Peptidase, M28D family.
 GN S03539.
 OS Shewanella oneidensis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
 OC Alteromonadaceae; Shewanella.
 NCBI_TaxID=70863;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MR-1;
 RA MEDLINE=22297686; PubMed=12368813;
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
 RA Read T.D., Eisen J.A., Seehadri R., Ward N., Methe B., Clayton R.A.,
 RA Meyer T., Tsaplin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
 RA Madupu R., Peterson R.J., Umayam L.A., White O., Wolf A.M.,
 RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
 RA Mueller J., Khouri H., Gill J., Ullrich T.R., McDonald L.A.,
 RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
 Shewanella oneidensis."
 RL Nat. Biotechnol. 20:1118-1123(2002).
 DR EMBL; AB015790; AAN56530.1;
 DR TIGR; SO3539;
 SQ SEQUENCE 414 AA; 44028 MW; C8AEB2A5773DBCB4 CRC64;

Query Match 20.4%; Score 180.5; DB 16; Length 414;
 Best Local Similarity 38.8%; Pred. No. 9,8e-09; Indels 1; Gaps 1;
 Matches 40; Conservative 17; Mismatches 45; Indels 1; Gaps 1;

Qy 74 VGPRLSSKNLEKAIQIMYQNLQODGLEKYLHPVIRPHMERGESAVMLEPRIRIKHAIIL 133
 Db 3 VGPRLSSKNLEKAIQIMYQNLQODGLEKYLHPVIRPHMERGESAVMLEPRIRIKHAIIL 105
 Qy 134 GLSSIGTPPEGITAEVLVVTSPFDELQRRASERARKIVVYQ 175
 Db 63 ALGGSVATPEGIKAKIARFNSLEALQATPDDVKGKIAFLDQ 105

RESULT 11

ID Q9A5A8 PRELIMINARY; PRT; 467 AA.
 AC Q9A5A8;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Aminopeptidase, putative.
 GN CC2544.
 OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173699; PubMed=11259647;
 RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N.D., Maddock J.R.,
 RA Ptočka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.D., Durkin A.S., Gilm M.L., Haft D.H.,
 RA Kolonay J.F., Sait J., Craven M.B., Khouli H., Shetty J., Berry K.,
 RA Kiehlbeck T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005922; AKZ4515.1; -.
 DR TIGR; CC2544; -.
 DR InterPro; IPR003137; PA.
 DR PROSITE; PS50840; PA; 1.
 KM Complete proteome.
 SQ SEQUENCE 467 AA; 4898 MW; 9984DE2A9D84F00F CRC64;
 Query Match 18.5%; Score 163.5; DB 16; Length 467;
 Best Local Similarity 35.5%; Pred. No. 4.8e-07;
 Matches 38; Conservative 17; Mismatches 51; Indels 1; Gaps 1;
 QY 70 LVDTGPRLSGSKNLEKAIQIWMYONLQODGLEKHLPEVPIPMWREGEASVAMLEPRH 129
 DB 49 LTTNIGRLVSGSPAMAKKMSVAKFKALGFTNINKEVFAKPSRGEESAEIWAAYAMK 108
 QY 130 IAILGSSIGTPPEGITAEVLVTSFDELQRRASEA-RGKIVYVNO 175
 DB 109 LGAVGLGRVSTPAGIGAEVALFKTRADMAAPDGLAKKIVITQ 155
 RESULT 12
 Q8PHE8 PRELIMINARY; PRT; 472 AA.
 AC O8PHE8;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Antinopeptidase.
 GN XAC3309.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
 RA Camarotte G., Canavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A.F., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinoia L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezra R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL; AE011976; AAM8152.1; -.
 DR InterPro; IPR003137; PA.
 KM Complete proteome.
 SQ SEQUENCE 375 AA; 38756 MW; 75F4A4914E2EA05 CRC64;

DR PROSITE; PS50840; PA; 1.
 KM Complete proteome.
 SQ SEQUENCE 472 AA; 48788 MW; 9956CDF1F49FC68 CRC64;
 Query Match 13.7%; Score 121; DB 16; Length 472;
 Best Local Similarity 27.0%; Pred. No. 0.0053;
 Matches 44; Conservative 20; Mismatches 75; Indels 24; Gaps 5;
 QY 12 VHLISCSGKAIKNGI---SKRTFEIKERISCGDVAKAIINLAVYGAQRYSYERLA 68
 DB 8 ISALLSCSCLAAQTSPDSALRTAAQLREQ-ALADTGFRAVQ----- 50
 QY 69 LVDTGPRLSGSKNLEKAIQIWMYONLQODGLEKHLPEVPIPMWREGEASVAMLEPRH 128
 DB 51 SLTTEVGPRRAGEADPRAVAMAKAFASIGPDKWTEPTPTPKMERSRSHAVIGHAQ 110
 QY 129 KIAIILGSSIGTPPEGITAEVLVTSFDELQRR-ASEARGKI 170
 DB 111 PLTITLGGSGPGGTVEG---EIVRFETLALQAAPAGSLAGKI 150
 RESULT 13
 Q8P625 PRELIMINARY; PRT; 375 AA.
 AC Q8P625;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Antinopeptidase.
 GN XCC3157.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
 RA Camarotte G., Canavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A.F., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinoia L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezra R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL; AE012431; AAM42427.1; -.
 DR InterPro; IPR003137; PA.
 DR PROSITE; PS50840; PA; 1.
 KM Complete proteome.
 SQ SEQUENCE 375 AA; 38756 MW; C75F4A4914E2EA05 CRC64;
 Query Match 13.3%; Score 117; DB 16; Length 375;
 Best Local Similarity 30.1%; Pred. No. 0.0094;
 Matches 34; Conservative 16; Mismatches 59; Indels 4; Gaps 2;
 QY 59 AONRSYERLALVDTPGPRLSGSKNLEKAIQIWMYONLQODGLEKHLPEVPIPMWREGE 118
 DB 41 ADSTGPAVVSLSLTTEVGPRRAGEADPRAVAMAKAFKISGIPDKWTEPTPTPKMERSE 100
 QY 119 SAVMLEPRHKAIIILGSSIGTPPEGITAEVLVTSFDELQRR-ASEARGKI 170
 DB 101 QAAVIGHAOPHLITLGGSGPGGTVEG---EIVRFETLALQAAPAGSLAGKI 150

RESULT 14

Q9WZ59 PRELIMINARY; PRT; 557 AA.
 ID Q9WZ59;
 AC Q9WZ59;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Pericillin-binding protein 2.
 GN TM0590.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogaceae; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Uitterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotten M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima."
 RL Nature 398:323-329 (1999).
 DR EMBL; AE001734; AAD35675.1; -.
 DR TIGR; TM0590; -.
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR005311; PBP_dimer.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF03717; PBP_dimer; 1.
 DR Pfam; PF09005; Transpeptidase; 1.
 DR PROSITE; PS00038; HLH_1; 1.
 DR Complete proteome.
 SQ SEQUENCE 557 AA; 61900 MW; DCBF394417124A80 CRC64;
 Query Match 10.6%; Score 93.5; DB 16; Length 557;
 Best Local Similarity 26.5%; Pred. No. 2.7; Mismatches 43; Conservative 26; Indels 51; Gaps 11;
 Matches 43; Conservative 26; Mismatches 43; Indels 51; Gaps 11;
 QY 20 GKAIICKNGIS-----KRTFERIK-----EIASCGDVAKAIINLA---VYGRK----- 59
 DB 58 GKVLAKDEVVYVLDPMWLSIDELKKTGLFTPEILSLVKGESIVDKARADVISKAGMRV 117
 QY 60 ---QNRSTYRLALVDVTGPRISGSKULEKA---IQIMYQNL---QDDGLEKVLHP--- 107
 DB 118 VMDYARKKTEPLA-----PHVGYVNADRVAGVGVESVYDFLGTGDKVKKVFEVPSGA 170
 QY 108 -----VRIPHWERGESAVMLEPRIRHKIAILGLGSSIGTPPEGI-----TAEVLVWTSFD 157
 DB 171 ISSEVLARSP-KPGEDVTLITDTRIQVAESL-EKIGNPGSVILSDVTRGSHIALASFP 228
 QY 158 E 158
 DB 229 E 229

RESULT 15

Q22257 PRELIMINARY; PRT; 1295 AA.
 ID Q22257;
 AC Q22257;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE T06E4.1 protein.
 GN T06E4.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Lloyd C.R.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018 (1998).
 DR EMBL; Z70756; CAA94789.1; -.
 DR WormPep; T06E4.1; CE06368.
 SQ SEQUENCE 1295 AA; 147395 MW; 57C70DBCCD2172F8 CRC64;

Query Match 10.5%; Score 92.5; DB 5; Length 1295;
 Best Local Similarity 20.4%; Pred. No. 10; Mismatches 33; Conservative 44; Indels 54; Gaps 5;
 Matches 33; Conservative 44; Mismatches 33; Indels 54; Gaps 5;
 QY 23 ICKNGISRTFEEIKERIASCGDVAKAIINLA---VYGRK-----DGLRLVDTLTKTKI---E 824
 DB 774 VMESEVLKESFEALQLELSASQEVSRSVDAAYOK-----DGLRLVDTLTKTKI---E 824
 QY 83 NLEKAIQIMYQNLQDDGLEKVLHPVRIPIHWERGESAVMLEPRIRHKIAILGLGSSIG 139
 DB 825 DTERKA---QDLOQSSVVEIKQLQDLQNFQKQVAEVLSEINEXLNSHHRDWMVALASQL 880
 QY 140 -----GTPPEGITAEVLVWTSFDELORRASEA 166
 DB 881 EELQHLKLVGESQVENYKEELIGAKINKEMKEMVDLNAKGLDA 922

Search completed: December 22, 2003, 11:51:54
 Job time : 28.507 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:49:10 ; Search time 14.0649, Seconds

(without alignments)
526.445 Million cell updates/sec

Title: US-09-745-763-36_COPY_1_175
Perfect score: 883
Sequence: 1 MKFLIFAFGCVHLLSLCSG.....FDELQRRASEARKIVVYVQ 175

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PTCUTS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	883	100.0	473	US-09-482-273-120	Sequence 120, App
2	82.5	9.3	337	US-09-107-532A-5998	Sequence 5998, Ap
3	81.5	9.2	262	US-09-328-352-4733	Sequence 4733, Ap
4	81	9.2	812	US-09-198-452A-978	Sequence 978, App
5	79.5	9.0	406	US-09-134-001C-3544	Sequence 3544, Ap
6	78.5	8.9	312	US-08-425-061-18	Sequence 18, Appl
7	78.5	8.9	312	US-08-825-886-18	Sequence 18, Appl
8	78.5	8.9	312	US-08-989-890-18	Sequence 18, Appl
9	78.5	8.9	765	US-08-425-061-19	Sequence 19, Appl
10	78.5	8.9	765	US-08-825-886-19	Sequence 19, Appl
11	78.5	8.9	765	US-08-989-890-19	Sequence 19, Appl
12	78.5	8.9	900	US-08-425-061-20	Sequence 20, Appl
13	78.5	8.9	900	US-08-825-886-20	Sequence 20, Appl
14	78.5	8.9	900	US-08-989-890-20	Sequence 20, Appl
15	78.5	8.9	914	US-08-425-061-21	Sequence 21, Appl
16	78.5	8.9	914	US-08-825-886-21	Sequence 21, Appl
17	78.5	8.9	914	US-08-989-890-21	Sequence 21, Appl
18	78.5	8.9	1202	US-08-425-061-22	Sequence 22, Appl
19	78.5	8.9	1202	US-08-825-886-22	Sequence 22, Appl
20	78.5	8.9	1202	US-08-989-890-22	Sequence 22, Appl
21	78.5	8.9	1363	US-08-425-061-23	Sequence 23, Appl
22	78.5	8.9	1363	US-08-825-886-23	Sequence 23, Appl
23	78.5	8.9	1363	US-08-989-890-23	Sequence 23, Appl
24	78.5	8.9	1852	US-08-425-061-24	Sequence 24, Appl
25	78.5	8.9	1852	US-08-825-886-24	Sequence 24, Appl
26	78.5	8.9	1852	US-08-989-890-24	Sequence 24, Appl
27	78.5	8.9	1863	US-08-425-061-16	Sequence 16, Appl

28	78.5	8.9	1863	1	US-08-598-591-2	Sequence 2, Appl
29	78.5	8.9	1863	1	US-08-480-784-2	Sequence 2, Appl
30	78.5	8.9	1863	1	US-08-483-553-2	Sequence 2, Appl
31	78.5	8.9	1863	1	US-08-487-002-2	Sequence 2, Appl
32	78.5	8.9	1863	1	US-08-483-554B-2	Sequence 2, Appl
33	78.5	8.9	1863	1	US-08-798-691-2	Sequence 2, Appl
34	78.5	8.9	1863	1	US-08-798-691-2	Sequence 2, Appl
35	78.5	8.9	1863	1	US-08-798-691-6	Sequence 6, Appl
36	78.5	8.9	1863	1	US-08-488-011B-2	Sequence 2, Appl
37	78.5	8.9	1863	2	US-08-825-886-16	Sequence 16, Appl
38	78.5	8.9	1863	2	US-08-603-753D-2	Sequence 2, Appl
39	78.5	8.9	1863	3	US-08-825-487A-2	Sequence 2, Appl
40	78.5	8.9	1863	3	US-08-825-487A-4	Sequence 4, Appl
41	78.5	8.9	1863	3	US-08-825-487A-6	Sequence 6, Appl
42	78.5	8.9	1863	3	US-09-074-476-2	Sequence 2, Appl
43	78.5	8.9	1863	3	US-09-074-476-4	Sequence 4, Appl
44	78.5	8.9	1863	3	US-09-074-476-6	Sequence 6, Appl
45	78.5	8.9	1863	3	US-09-099-753-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-482-273-120
; Sequence 120, Application US/09482273
; Patent No. 6534631
;
GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; EARLIER FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 120
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (473)
; OTHER INFORMATION: Xaa equals stop translation
;
US-09-482-273-120
Query Match 100.0%; Score 883; DB 4; Length 473;
Best Local Similarity 100.0%; Pred. No. 1.1e-93;
Matches 175; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 MKFLIFAFGCVHLLSLCSGKALCKNGISKRTPEIRIEETASCGDVAKAIINLAVYGAQ 60
DB 1 MKFLIFAFGCVHLLSLCSGKALCKNGISKRTPEIRIEETASCGDVAKAIINLAVYGAQ 60
QY 61 NRSYERLALVDVTPGPRLSGSKNLEKAIQIMYQNLQDDGLEKYLEPVRIPHWERGESA 120
DB 61 NRSYERLALVDVTPGPRLSGSKNLEKAIQIMYQNLQDDGLEKYLEPVRIPHWERGESA 120
QY 121 VMLPRLHKLIALILGLSSIGTPPEGITAEVLVTVSPDELQRRASEARKIVVYVQ 175
DB 121 VMLPRLHKLIALILGLSSIGTPPEGITAEVLVTVSPDELQRRASEARKIVVYVQ 175
RESULT 2
US-09-107-532A-5998
; Sequence 5998, Application US/09107532A
```

Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5998:
SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (8) LOCATION 1...337
SEQUENCE DESCRIPTION: SEQ ID NO: 5998:
US-09-107-532A-5998
Query Match
Best Local Similarity 9.3%; Score 82.5; DB 4; Length 337;
Matches 43; Conservative 30; Mismatches 44; Indels 99; Gaps 9;
QY 7 AFGGVHLISLCSGKAIKNGISKRTFEIKET-----ASGSDVAKAIIN 52
DB 2 SLGGIIRKNN-----KDLVAKAVYDVVKDLDLLEQVEQLLENPKSEHSDVAFPAFS 53
QY 53 LA-VGSKAQNREYERLALLVDT-----VGPRLSGSKNLE---KAIQIMYONLQODG 99
DB 54 LAKVTRKAPQADLADEKIDISANFEKIEVVGPIYINFPFNKKELISKVQLQTVVKEKEHYG 113
QY 100 -----LEKVALEPRIPH---W--- 113
DB 114 DSNIGNOGTVPIIDMSPNIAKPIISMGHLRSTVIGNSIGRIMKIGVQPIRINHLDWGTQ 173
QY 114 -----ERGESAVMLEP-----RIHKIA 131
DB 174 FGKLIIVAYKKWGTBEAVVAKBPINELLRLVYQFHEVA 209

RESULT 3
US-09-328-352-4733
Sequence 4733, Application US/09328352
Patent No. 6562958

GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 4733
LENGTH: 262
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-4733
Query Match
Best Local Similarity 9.2%; Score 81.5; DB 4; Length 262;
Matches 31; Conservative 20; Mismatches 34; Indels 11; Gaps 5;
QY 67 LALLVDTVGPRLSGSKNLEKAIQIMYONLQODGLEKVALEPRIP---HMERGESAVML 123
DB 85 LSTIVQTRTP-LDGRFALVALANI-----LQIPQLPISLQ-VKMPNDLYSTGKMGGLIV 138
QY 124 EPRHKIALLGLSSIGTPPEGITAEVLVVTSPDEL 159
DB 139 EPLSQHQAIVGVGINLKTTP--VTESDQPISTLELD 172
RESULT 4
US-09-198-452A-978
Sequence 978, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 978
LENGTH: 812
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-978
Query Match
Best Local Similarity 9.2%; Score 81; DB 4; Length 812;
Matches 48; Conservative 28; Mismatches 62; Indels 60; Gaps 9;
QY 17 LCGSKAIKNGISKRTFEIKETIA--SCGVAKAIINL-----AVYKAQ 60
DB 460 VCGGKS-CEHDISLSAQHVSQVISPFFYDVSYILNRQGLMRTGKOPHILIEGTQDSP 518
QY 61 NRSYERLALL-VDTVGPRLSGSKNLEKAIQIMYONLQOD----- 98
DB 519 LSSRIASLAKVDDLFPVLHGFEGEDGTIGFFELIKPVAQPSLSLAATAMDKLTKRI 578
QY 99 -----GLEKVALEPRIPHERGESAV--MLE---PRIHKIALLGLSSIGTPPEGITA 148
DB 579 ASAVGVVVPVQPIINLTFWKRNPCLQINLIETSPFMIVXTA--HIGSSIG----- 628
QY 149 EVLVVTSFDELQRASEA 166
DB 629 -IFLVKDELEQEKISA 645
RESULT 5
US-09-134-001C-3544
Sequence 3544, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;; FILE REFERENCE: GTC-007
;; CURRENT APPLICATION NUMBER: US/09/134,001C
;; PRIOR FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 3544
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3544

Query Match 9.0%; Score 79.5; DB 4; Length 406;
Best Local Similarity 24.0%; Pred. No. 0.97;
Matches 44; Conservative 34; Mismatches 58; Indels 47; Gaps 9;

QY 3 FLIFAFGSHLS--LCSGKAIKNGISKRTFEIEKEIASCGDAKAI-----NL 53
DB 87 FLIDIKTYSHLSRLVSGNSV-----VYQLEQALSEHPFEDALFNSGDANL 138
QY 54 AVYGAQNRS-----YERLALVDVTPRLSGSKNLEKAIQIWMYONLQODGLEKVALEP 107
DB 139 AVFIIFKNNNVIVISDQNHASIID--GIKLSGS-----KVIYHLNDLDES-HLAR 189
QY 108 VRIPIHMER-----GEES---AVMLEPRIHKAIIIGLSSIGTPEGITAEVLV 152
DB 190 HTNPDVQKIVISDVSTNGTKADINRLVHLKQRYNAIILIDASHSLGMLFEYHADIDI 249
QY 153 VTS 155
DB 250 VTS 252

RESULT 6

US-08-425-061-18
; Sequence 18, Application US/08425061
; Patent No. 5622829
; GENERAL INFORMATION:
; APPLICANT: KING, Mary-Claire
; APPLICANT: FRIEDMAN, Lori
; APPLICANT: OSTERMEYER, Beth
; APPLICANT: ROWELL, Sarah
; APPLICANT: LYNCH, Eric
; APPLICANT: SZABO, Csilla
; APPLICANT: LEE, Ming
; TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN
; TITLE OF INVENTION: CANCER
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,061
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO
; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (415) 494-8700
;; TELEFAX: (415) 494-8771
;; TELEX: 910 277299
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 312 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-425-061-18

Query Match 8.9%; Score 78.5; DB 1; Length 312;
Best Local Similarity 22.8%; Pred. No. 0.85;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLSLCSGKA---IKNGISKRT-----FEEIEKEIAS--CG--DVAKAI 50
DB 42 IFCKFCMLKLNQKKPSQCLCKNDITKRSIQESTRFSLVLEIKIICAFOLDTGLEY 101
QY 51 INLAIVGKQNRSEYERL---ALLVDVTPRLSGSKNL-----EKAQIWMYONL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSTIQSMGYRNRKRLQSEPPNSIQETISLSVOLSNLGT 161
QY 96 QODGLEKVALEPRIP-HMERGESAVMLEPRIHKAIIIGLSS--IGTPEGITAEVLV 152
DB 162 VRLIRTKRQIQPKTSVYIELSSDS---EDTVNKAITYGSGQDELLQITPGTBEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKAACE 227

RESULT 7

US-08-825-886-18
; Sequence 18, Application US/08825886
; Patent No. 5821328
; GENERAL INFORMATION:
; APPLICANT: KING, Mary-Claire
; APPLICANT: FRIEDMAN, Lori
; APPLICANT: OSTERMEYER, Beth
; APPLICANT: ROWELL, Sarah
; APPLICANT: LYNCH, Eric
; APPLICANT: SZABO, Csilla
; APPLICANT: LEE, Ming
; TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN
; TITLE OF INVENTION: CANCER
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/825,886
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/425,061
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids.
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-825-886-18

Query Match 8.9%; Score 78.5; DB 2; Length 312;
Best Local Similarity 22.8%; Pred. No. 0.85;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGGVHLISLCSGKA---ICKNGISKRT-----FEETKEETIAS--CG---DVAKAI 50
DB 42 IFCKFCMLKLLNQKKGPSOCPKCDITKRSLOESTRFSQVLEELKTIICAFQDGTGL 101
QY 51 INLAVYGAQNRSYERL---ALLVDTVGPRLSGSKNL-----EKAIQIMYQNL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNARAKLLQSEPEPNSLOETSLVSQLSNLGT 161
QY 96 QODGLEKXVHLEPVRI-P-HMERGESAVMLEPPIHKIAILGLGSS--IGTPPEGITAEVLV 152
DB 162 VRTIARTKRIOPKRTSVYIELGSDSS---EDTVNKATYCSVGDQELQITPGTRDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKAACE 227

RESULT 8

US-08-989-890-18

Sequence 18, Application US/08989890
Patent No. 6512091

GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire

FRIEDMAN, Lori

OSTERMEYER, Beth

ROWELL, Sarah

LYNCH, Eric

SZABO, Csilla

LEE, Ming

TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESS: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,890

FILING DATE: 12-Dec-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/825,886

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/425,061

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8771

TELEFAX: 910 277299

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 312 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-08-989-890-18

Query Match 8.9%; Score 78.5; DB 4; Length 312;
Best Local Similarity 22.8%; Pred. No. 0.85;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGGVHLISLCSGKA---ICKNGISKRT-----FEETKEETIAS--CG---DVAKAI 50
DB 42 IFCKFCMLKLLNQKKGPSOCPKCDITKRSLOESTRFSQVLEELKTIICAFQDGTGL 101
QY 51 INLAVYGAQNRSYERL---ALLVDTVGPRLSGSKNL-----EKAIQIMYQNL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNARAKLLQSEPEPNSLOETSLVSQLSNLGT 161
QY 96 QODGLEKXVHLEPVRI-P-HMERGESAVMLEPPIHKIAILGLGSS--IGTPPEGITAEVLV 152
DB 162 VRTIARTKRIOPKRTSVYIELGSDSS---EDTVNKATYCSVGDQELQITPGTRDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKAACE 227

RESULT 9

US-08-425-061-19

Sequence 19, Application US/08425061
Patent No. 5622829

GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire

FRIEDMAN, Lori

OSTERMEYER, Beth

ROWELL, Sarah

LYNCH, Eric

SZABO, Csilla

LEE, Ming

TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESS: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,061

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 765 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-425-061-19

Query Match 8.9%; Score 78.5; DB 1; Length 765;
Best Local Similarity 22.8%; Pred. No. 3.2;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLCSGKA---ICKNGISKRT-----FEEIKEIAS--CG---DVAKAI 50
DB 42 IFKFCMLKILNKKGSGQCLCKNDITKSLQESTRFSGVLELLKIIICAFQDITGLV 101
QY 51 INLAIVGKAONRYSERL---ALVDVGPRLSSKUL-----EKAIQIMYONL-- 95
DB 102 ANSYNPAKKENNSPEHLKDEVSIIQSWGYNRAKRLQSEPNPSIQETSLVQLSNLGT 161
QY 96 QODGLEKVLHPVRI-P-HMERGESAVMLEPRHIIKIALIGLSS--IGTPEGITAVLV 152
DB 162 VRLIRTKQRIQPQTSVYIELGSDS---EDVINKATYCSVGQDELLQITPQGTREI-- 216
QY 153 VTSFDELORRASE 165
DB 217 --SLDSAKKAACE 227

RESULT 10

US-08-825-886-19
Sequence 19, Application US/08825886
Patent No. 5823328

GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire
APPLICANT: FRIEDMAN, Lori
APPLICANT: OSTERMEYER, Beth
APPLICANT: ROWELL, Sarah
APPLICANT: LYNCH, Eric
APPLICANT: SZABO, Csilla
APPLICANT: LEE, Ming
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN
TITLE OF INVENTION: CANCER
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/825, 886
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/425, 061
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
LENGTH: 765 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-825-886-19

Query Match 8.9%; Score 78.5; DB 2; Length 765;
Best Local Similarity 22.8%; Pred. No. 3.2;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLCSGKA---ICKNGISKRT-----FEEIKEIAS--CG---DVAKAI 50
DB 42 IFKFCMLKILNKKGSGQCLCKNDITKSLQESTRFSGVLELLKIIICAFQDITGLV 101
QY 51 INLAIVGKAONRYSERL---ALVDVGPRLSSKUL-----EKAIQIMYONL-- 95
DB 102 ANSYNPAKKENNSPEHLKDEVSIIQSWGYNRAKRLQSEPNPSIQETSLVQLSNLGT 161
QY 96 QODGLEKVLHPVRI-P-HMERGESAVMLEPRHIIKIALIGLSS--IGTPEGITAVLV 152
DB 162 VRLIRTKQRIQPQTSVYIELGSDS---EDVINKATYCSVGQDELLQITPQGTREI-- 216
QY 153 VTSFDELORRASE 165
DB 217 --SLDSAKKAACE 227

RESULT 11

US-08-989-890-19
Sequence 19, Application US/08989890
Patent No. 6512091

GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire
APPLICANT: FRIEDMAN, Lori
APPLICANT: OSTERMEYER, Beth
APPLICANT: ROWELL, Sarah
APPLICANT: LYNCH, Eric
APPLICANT: SZABO, Csilla
APPLICANT: LEE, Ming
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN
TITLE OF INVENTION: CANCER
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989, 890
FILING DATE: 12-Dec-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/825, 886
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/425, 061
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 765 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-08-989-890-19

Query Match 8.9%; Score 78.5; DB 4; Length 765;
Best Local Similarity 22.8%; Pred. No. 3.2;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLISLCSGKA---ICKNGISKRT-----PEEIKERIAS--CG---DVAKAI 50
DB 42 IFCKFCMKLTKNQKGPSCPLCKNDITKRSLOESTRFSQVLEIKTICAFQDITGLEV 101
QY 51 INLAVYGAONRSYERL---ALLVDTVGPRLSSKNL-----EKAIQIMYQNL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNRAKRLLOSEPENPSLOETSLSVOLSNLGT 161
QY 96 QODGLEKXVLEPVRI-P-HMERGESAVMLEPRIHKAIIILIGSS--IGTPEEGITAEVLV 152
DB 162 VRLTTRKORIQPKTSVYIELGSDSS---EDTVNKATYCSVGDELLQITPGTRDEI-- 216
QY 153 VTSFDELQRPASE 165
DB 217 --SLDSAKKAACE 227

RESULT 12
US-08-425-061-20
Sequence 20, Application US/08425061
Patent No. 5622829
GENERAL INFORMATION:
APPLICANT: KING, Mary-Claire
APPLICANT: FRIEDMAN, Lori
APPLICANT: OSTERMEYER, Beth
APPLICANT: ROWELL, Sarah
APPLICANT: LYNCH, Eric
APPLICANT: SZABO, Csilla
APPLICANT: LEE, Ming
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN
TITLE OF INVENTION: CANCER
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425.061
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 900 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-425-061-20

Query Match 8.9%; Score 78.5; DB 1; Length 900;
Best Local Similarity 22.8%; Pred. No. 4.1;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLISLCSGKA---ICKNGISKRT-----PEEIKERIAS--CG---DVAKAI 50
DB 42 IFCKFCMKLTKNQKGPSCPLCKNDITKRSLOESTRFSQVLEIKTICAFQDITGLEV 101
QY 51 INLAVYGAONRSYERL---ALLVDTVGPRLSSKNL-----EKAIQIMYQNL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNRAKRLLOSEPENPSLOETSLSVOLSNLGT 161
QY 96 QODGLEKXVLEPVRI-P-HMERGESAVMLEPRIHKAIIILIGSS--IGTPEEGITAEVLV 152
DB 162 VRLTTRKORIQPKTSVYIELGSDSS---EDTVNKATYCSVGDELLQITPGTRDEI-- 216
QY 153 VTSFDELQRPASE 165
DB 217 --SLDSAKKAACE 227

RESULT 13
US-08-825-886-20
Sequence 20, Application US/08825886
Patent No. 5821328
GENERAL INFORMATION:
APPLICANT: KING, Mary-Claire
APPLICANT: FRIEDMAN, Lori
APPLICANT: OSTERMEYER, Beth
APPLICANT: ROWELL, Sarah
APPLICANT: LYNCH, Eric
APPLICANT: SZABO, Csilla
APPLICANT: LEE, Ming
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN
TITLE OF INVENTION: CANCER
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/825.886
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/425,061
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 900 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-825-886-20

Query Match 8.9%; Score 78.5; DB 2; Length 900;

Best Local Similarity 22.8%; Pred. No. 4.1;

Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLCSGKA---ICKNGISKRT-----FEEIKERIAS--CG--DVAKAI 50
DB 42 IFCFCMKILNQKKGSPCLCKNDITKSLQSTRFSLQVEELKIKICAFQIDTGLEV 101
QY 51 INLAVYKKAQNRSYERL---ALVDYVGPRLSGSKNL-----EKAIQMYQNL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNARAKRLQSEPPENSLQETSLSVQSLNIGT 161
QY 96 QQDGLKVHLEPVRIIP-HMERGESAVMLEPRHKIAILGIGSS--IGTPPEGITAEVLV 152
DB 162 VRTLRTRQRIQPKQTSYIHLGSDS---EDTVNKATYCSVGDELLQITPQGRDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKAACE 227

RESULT 14

US-08-989-890-20

Sequence 20, Application US/08989890

Patent No. 6512091

GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire

FRIEDMAN, Lori

OSTERMAYER, Beth

ROWELL, Sarah

LYNCH, Eric

SZABO, Ceilia

LEE, Ming

TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN

CANCER

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,890

FILING DATE: 12-Dec-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/825,886

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/425,061

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 900 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-08-989-890-20

Query Match 8.9%; Score 78.5; DB 4; Length 900;

Best Local Similarity 22.8%; Pred. No. 4.1;

Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLCSGKA---ICKNGISKRT-----FEEIKERIAS--CG--DVAKAI 50
DB 42 IFCFCMKILNQKKGSPCLCKNDITKSLQSTRFSLQVEELKIKICAFQIDTGLEV 101
QY 51 INLAVYKKAQNRSYERL---ALVDYVGPRLSGSKNL-----EKAIQMYQNL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNARAKRLQSEPPENSLQETSLSVQSLNIGT 161
QY 96 QQDGLKVHLEPVRIIP-HMERGESAVMLEPRHKIAILGIGSS--IGTPPEGITAEVLV 152
DB 162 VRTLRTRQRIQPKQTSYIHLGSDS---EDTVNKATYCSVGDELLQITPQGRDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKAACE 227

RESULT 15

US-08-425-061-21

Sequence 21, Application US/08425061

Patent No. 5622829

GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire

FRIEDMAN, Lori

OSTERMAYER, Beth

ROWELL, Sarah

LYNCH, Eric

SZABO, Ceilia

APPLICANT: LEE, Ming

TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN

CANCER

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,061

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 914 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

Tue Dec 23 11:24:39 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:47:05 ; Search time 14.6059 Seconds
(without alignments)
2237.645 Million cell updates/sec

Title: US-09-745-763-36_COPY_1_175
Perfect score: 883
Sequence: 1 MKRLIFAFGFGVHLISCSG.....FDEIQRASRARKIVYVQ 175

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Published Applications_AA:*
2: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/CTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	883	100.0	472	9 US-09-745-763-36	Sequence 36, Appl
2	883	100.0	473	11 US-09-984-771-120	Sequence 120, App
3	92.5	10.5	325	15 US-10-156-761-11911	Sequence 11911, A
4	81.5	9.2	162	12 US-09-882-227-594	Sequence 594, App
5	81	9.2	340	15 US-10-186-886-2	Sequence 2, Appl
6	79	8.3	1390	15 US-10-087-340-35	Sequence 35, Appl
7	78.5	8.9	1863	10 US-09-734-672-2	Sequence 2, Appl
8	78.5	8.9	1863	10 US-09-734-672-4	Sequence 4, Appl
9	78.5	8.9	1863	10 US-09-734-672-6	Sequence 6, Appl
10	78.5	8.9	1863	11 US-09-982-828-2	Sequence 2, Appl
11	78.5	8.9	1863	11 US-09-982-828-4	Sequence 4, Appl
12	78.5	8.9	1863	11 US-09-982-828-6	Sequence 6, Appl
13	78.5	8.9	1863	15 US-10-022-819-2	Sequence 2, Appl
14	74	8.4	384	9 US-09-925-299-1011	Sequence 1011, Ap
15	74	8.4	384	11 US-09-925-299-1011	Sequence 1011, Ap

16	74	8.4	390	8 US-08-731-566-2	Sequence 2, Appl
17	74	8.4	390	11 US-09-823-187-29	Sequence 29, Appl
18	74	8.4	390	11 US-09-823-187-30	Sequence 30, Appl
19	74	8.4	390	12 US-10-094-944-12	Sequence 12, Appl
20	74	8.4	619	11 US-09-036-614A-1	Sequence 1, Appl
21	74	8.4	619	12 US-10-094-749-1964	Sequence 1964, Ap
22	73.5	8.3	159	12 US-09-769-736-44	Sequence 44, Appl
23	73.5	8.3	243	12 US-10-225-068-182	Sequence 182, App
24	73.5	8.3	459	10 US-09-950-772-4	Sequence 4, Appl
25	73.5	8.3	459	15 US-10-289-360-2	Sequence 2, Appl
26	73.5	8.3	1565	12 US-10-262-794A-59	Sequence 59, Appl
27	73.5	8.3	1565	15 US-10-242-056-59	Sequence 59, Appl
28	73	8.3	848	9 US-09-925-301-1161	Sequence 1161, Ap
29	73	8.3	2353	10 US-09-797-862-33	Sequence 33, Appl
30	72.5	8.2	334	9 US-09-159-469-28	Sequence 28, Appl
31	72.5	8.2	334	9 US-09-159-469-73	Sequence 28, Appl
32	72.5	8.2	530	9 US-09-159-469-73	Sequence 73, Appl
33	72.5	8.2	530	9 US-09-798-042-73	Sequence 64, Appl
34	72.5	8.2	530	9 US-09-159-469-64	Sequence 64, Appl
35	72.5	8.2	530	9 US-09-798-042-64	Sequence 13744, A
36	72	8.2	478	15 US-10-156-761-13744	Sequence 2, Appl
37	72	8.2	1237	12 US-10-132-134-2	Sequence 6926, Ap
38	71.5	8.1	223	10 US-09-738-626-6926	Sequence 33, Appl
39	71.5	8.1	435	10 US-09-866-582-33	Sequence 1524, Ap
40	71.5	8.1	519	12 US-10-012-697-1524	Sequence 2, Appl
41	71.5	8.1	600	9 US-09-975-901-2	Sequence 209, App
42	71.5	8.1	1288	11 US-09-919-039-209	Sequence 131, App
43	71	8.0	486	12 US-10-381-779-131	Sequence 2525, Ap
44	71	8.0	619	12 US-10-094-749-2525	Sequence 6, Appl
45	71	8.0	2066	11 US-09-978-244A-6	

ALIGNMENTS

RESULT 1
US-09-745-763-36
Sequence 36, Application US/09745763
Patent No. US20020065394A1

GENERAL INFORMATION:
APPLICANT: McCosy, Kenneth
McCoy, John M.
LaValle, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Werberg, David
Treacy, Maurice
Spaulding, Vikki

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

```

; INFORMATION FOR SEQ ID NO: 36:
;   SEQUENCE CHARACTERISTICS:
;       LENGTH: 472 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: <Unknown>
;       TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-745-763-36

Query Match      100.0%; Score 883; DB 9; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.3e-91;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKELFAFGVGHLLSLGSKAICNGISKRTFEETKEEIASCGVAATINLAVYGAQ 60
DB 1 MKELFAFGVGHLLSLGSKAICNGISKRTFEETKEEIASCGVAATINLAVYGAQ 60
QY 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGES 120
DB 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGES 120
QY 121 VMLEPRIRHKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175
DB 121 VMLEPRIRHKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175

RESULT 2
US-09-984-271-120
; Sequence 120, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: 71 Human Secreted Proteins
;   FILE REFERENCE: P2030P1
;   CURRENT APPLICATION NUMBER: US/09/984,271
;   CURRENT FILING DATE: 2001-10-29
;   PRIOR APPLICATION NUMBER: 09/482,273
;   PRIOR FILING DATE: 2000-01-13
;   PRIOR APPLICATION NUMBER: PCT/US99/15849
;   PRIOR FILING DATE: 1999-07-14
;   PRIOR APPLICATION NUMBER: 60/092,921
;   PRIOR FILING DATE: 1998-07-15
;   PRIOR APPLICATION NUMBER: 60/092,922
;   PRIOR FILING DATE: 1998-07-15
;   PRIOR APPLICATION NUMBER: 60/092,956
;   PRIOR FILING DATE: 1998-07-15
;   NUMBER OF SEQ ID NOS: 267
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO: 120
;   LENGTH: 473
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (473)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-120

Query Match      100.0%; Score 883; DB 11; Length 473;
Best Local Similarity 100.0%; Pred. No. 4.3e-91;
Matches 175; Conservative 0; Mismatches .0; Indels 0; Gaps 0;

QY 1 MKELFAFGVGHLLSLGSKAICNGISKRTFEETKEEIASCGVAATINLAVYGAQ 60
DB 1 MKELFAFGVGHLLSLGSKAICNGISKRTFEETKEEIASCGVAATINLAVYGAQ 60
QY 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGES 120
DB 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGES 120
QY 121 VMLEPRIRHKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175
DB 121 VMLEPRIRHKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175
```

```

DB 121 VMLEPRIRHKIALILGSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175

RESULT 3
US-10-156-761-11911
; Sequence 11911, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;   APPLICANT: OMIDA, SATOSHI
;   APPLICANT: IKEDA, HARUO
;   APPLICANT: ISHIKAWA, JUN
;   APPLICANT: HORIKAWA, HIROSHI
;   APPLICANT: SHIBA, TADAYOSHI
;   APPLICANT: SAKAKI, YOSHIYUKI
;   APPLICANT: HATTORI, MASAHIRA
;   TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;   FILE REFERENCE: 249-262
;   CURRENT APPLICATION NUMBER: US/10/156,761
;   CURRENT FILING DATE: 2002-05-29
;   PRIOR APPLICATION NUMBER: JP 2001-204089
;   PRIOR FILING DATE: 2001-05-30
;   PRIOR APPLICATION NUMBER: JP 2001-272697
;   PRIOR FILING DATE: 2001-08-02
;   NUMBER OF SEQ ID NOS: 15109
;   SEQ ID NO: 11911
;   LENGTH: 325
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-11911

Query Match      10.5%; Score 92.5; DB 15; Length 325;
Best Local Similarity 21.5%; Pred. No. 0.077;
Matches 37; Conservative 29; Mismatches 73; Indels 33; Gaps 5;

QY 33 FEETKEEIASCGVAATINLAV-----YGAONRSYERLALVDVTGPRLSGSKN 83
DB 92 YDQIVTQLAKMKAARALKIKLPVVVRIPYGGIGAVHSESPALPAHVAAGLKVSPSN 151
QY 84 LEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESAVMLEPRIRHKIALILGSSI---- 139
DB 152 ASDAYVMWQQAIOISDD-PVIFPEPKR-RYWDKGSVNEALPDPILHKARVVRGGDTLLAA 209
QY 140 -----GTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 173
DB 210 YGPRVVKYCGEAAAAAEBSGKSLVVDLRSMSPIDFVAQASVETRLVVVH 261

RESULT 4
US-09-882-227-594
; Sequence 594, Application US/09882227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
;   APPLICANT: Kleantous, Harold
;   APPLICANT: Al-Garawi, Amal
;   APPLICANT: Miller, Charles
;   APPLICANT: Tomb, Jean-Francois
;   APPLICANT: Oocmen, Raymond P.
;   TITLE OF INVENTION: Identification of Polynucleotides
;   TITLE OF INVENTION: Encoding No. US20030158396A1 Helicobacter Polypeptides in the
;   FILE REFERENCE: 06132/047002
;   CURRENT APPLICATION NUMBER: US/09/882,227
;   CURRENT FILING DATE: 2001-06-15
;   PRIOR APPLICATION NUMBER: US 08/902,615
;   PRIOR FILING DATE: 1997-07-29
;   NUMBER OF SEQ ID NOS: 638
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO: 594
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-09-882-227-594
```

```
Query Match          9.2%; Score 81.5; DB 12; Length 162;
Best Local Similarity 23.5%; Pred. No. 0.49;
Matches 35; Conservative 31; Mismatches 60; Indels 23; Gaps 6;

Qy 28 ISKTFEIEIEEIASCGDVAKAIIINLAVYGAQNRSYERLALLVDTYGPRLSGSKNLEKA 87
Db 20 IGSFTKENIK-EYAKC-----FMSLSMFLEOEKNOOEFELEODT-----KENOEEL 66

Qy 88 IQIYQNLQODGLEKHLPEVRIPIHMERGESAVMLBRHKIKAI-----LGLSSSIGPP 143
Db 67 I---KNIQSIANKOELEKISFERWENKIOERVL--PKLRIVTHKQESITSSINTQL 120

Qy 144 EGITAEVLVYTSFDELQPRASEARKIIV 172
Db 121 ESFKDELDSVFEIQKNTQIAYRLAI 149

RESULT 5
US-10-186-886-2
; Sequence 2, Application US/10186886
; Publication No. US20030119061A1
; GENERAL INFORMATION:
; APPLICANT: Navia, Manuel A.
; APPLICANT: Ala, Paul J.
; APPLICANT: Griffith, James P.
; APPLICANT: Ali, Janid A.
; APPLICANT: Faerman, Carlos H.
; APPLICANT: Moe, Scott T.
; APPLICANT: Magee, Andrew S.
; APPLICANT: Connelly, Patrick R.
; APPLICANT: Perola, Emanuele
; TITLE OF INVENTION: STRUCTURE-BASED DRUG DESIGN METHODS FOR
; TITLE OF INVENTION: IDENTIFYING D-ALA-D-ALA LIGASE INHIBITORS AS ANTI-BACTERIAL
; FILE REFERENCE: 10283-014001
; CURRENT APPLICATION NUMBER: US/10/186,886
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 60/301,676
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Chlamydomophila pneumoniae
US-10-186-886-2

Query Match          9.2%; Score 81; DB 15; Length 340;
Best Local Similarity 24.2%; Pred. No. 1.6;
Matches 48; Conservative 28; Mismatches 62; Indels 60; Gaps 9;

Qy 17 LSCGKAICNGISKTFEIEIEEIASCGDVAKAIIIN-----AVYGAQ 60
Db 9 VCGSKS-CEHDSISLSAQHSKYISPEFYDVSFYIINQGLMRGKDPHLEETOGDSP 67

Qy 61 NRSYERLALL-VDTYGPRLSGSKNLEKAIQIMYQNLQOD----- 98
Db 68 LSSEIASLAKVDCLEFVLHGPFGEDGTIQGFELLGKPVAGPSISLATAVMDKLLTKRI 127

Qy 99 -----GLEKVLHPEVRIPIHMERGESAV--MLE-----PRIHKIAIILGLSSSIGTPEGITTA 148
Db 128 ASAVGVPPVYQPLNLCFKWKNPELCTQNLITETFPPIVAKTA--HGLSSSIG----- 177

Qy 149 EVLVVTSFDELQPRASEA 166
Db 178 -IFLVDRKEELQEKISEA 194

RESULT 6
US-10-097-340-35
; Sequence 35, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
```

```
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAPARAPU
APPLICANT: Sebastian HOERSCH
APPLICANT: Shubhangi KAMATYAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 1390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-35

Query Match          8.9%; Score 79; DB 15; Length 1390;
Best Local Similarity 28.2%; Pred. No. 21;
Matches 37; Conservative 26; Mismatches 42; Indels 26; Gaps 9;

Qy 15 LSLCGKAICNGISKTFEIEIEEIASCGDVAKAIIIN---LAVYGAQNRSYERLALLV 71
Db 851 LSLLE-RVLMKDIYTPPOSEFVKTVIRKCLEQA-ALVNYGRLSYAKIEENQD-----A 903

Qy 72 DTVGPRLSGSKNLEKAI--QIMYQNLQODGLEKHLPEVRIPIHMERGEE---SAVMLE 124
Db 904 ENVGRLITPAKKLEDITRLAEVLVLEVQON--EEHHAEE---PVDGGEAFAMWSDIME 957

Qy 125 PRIHKIAIILGL 135
Db 958 ---HAETFLSL 965

RESULT 7
US-09-734-672-2
; Sequence 2, Application US/09734672
; Publication No. US20020183268A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Patricia D.
; APPLICANT: Allen, Antonette C.
; APPLICANT: Alvarez, Christopher P.
; APPLICANT: Critz, Brenda S.
; APPLICANT: Olson, Sheri J.
; APPLICANT: Schelter, Denise B.
; APPLICANT: Zeng, Bin
; TITLE OF INVENTION: Coding Sequences of the Human
```

BCR1 Gene
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan Lewis & Bockius LLP
STREET: 1111 Pennsylvania Ave., N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/734,672
FILING DATE: 03-Dec-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/966,436
FILING DATE: 07-No. US20020183268A1-97
APPLICATION NUMBER: US 08/598,591
FILING DATE: 12-Feb-96
ATTORNEY/AGENT INFORMATION:
NAME: Michael S. Tuscan
REGISTRATION NUMBER: 43,210
REFERENCE/DOCKET NUMBER: 44921-5055-02-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-739-3000
TELEFAX: 202-739-3001
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020183268A1 Relevant
TOPOLOGY: No. US20020183268A1 Relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: BRC1
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 17
MAP POSITION: 17q21
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-734-672-2

Query Match 8.9%; Score 78.5; DB 10; Length 1863;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLISLSCGKA---ICKNGISKRT-----FEEIKERIAS--CG---DVAKAI 50
DB 42 IFCKCMKLNQKKGSCPCNDITKSLQESTRFSQVBLELKIICAFQDITGLV 101
QY 51 INLAVYGAQNSYERL---ALLVDTGPRISGSKNL-----EKAIQIYQNL-- 95
DB 102 ANSYNPAKKENSPHLKDEVSIIQSMGYRNRAKRLQSEPNPSLQETSLSVQSLNIGT 161
QY 96 QODGLEKHLPEVRIP-HMERGESAVMLEPRIKIAILGLSS--IGTPEGITAEVLV 152
DB 162 VRTLRTKRIOPQKTSYVIELGSDS---EDTVNKATYCSVGDQLQITPGTRDEI-- 216
QY 153 VTSFDELORASE 165
DB 217 --SLDSAKKACE 227

RESULT 8
US-09-734-672-4
Sequence 4. Application US/09734672
Publication No. US20020183268A1
GENERAL INFORMATION:
APPLICANT: Murphy, Patricia D.
Allen, Antonette C.

Alvarez, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Schelter, Denise B.
Zeng, Bin
TITLE OF INVENTION: Coding Sequences of the Human
BCR1 Gene
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan Lewis & Bockius LLP
STREET: 1111 Pennsylvania Ave., N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/734,672
FILING DATE: 03-Dec-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/966,436
FILING DATE: 07-No. US20020183268A1-97
APPLICATION NUMBER: US 08/598,591
FILING DATE: 12-Feb-96
ATTORNEY/AGENT INFORMATION:
NAME: Michael S. Tuscan
REGISTRATION NUMBER: 43,210
REFERENCE/DOCKET NUMBER: 44921-5055-02-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-739-3000
TELEFAX: 202-739-3001
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020183268A1 Relevant
TOPOLOGY: No. US20020183268A1 Relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: BRC1
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 17
MAP POSITION: 17q21
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-734-672-4

Query Match 8.9%; Score 78.5; DB 10; Length 1863;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLISLSCGKA---ICKNGISKRT-----FEEIKERIAS--CG---DVAKAI 50
DB 42 IFCKCMKLNQKKGSCPCNDITKSLQESTRFSQVBLELKIICAFQDITGLV 101
QY 51 INLAVYGAQNSYERL---ALLVDTGPRISGSKNL-----EKAIQIYQNL-- 95
DB 102 ANSYNPAKKENSPHLKDEVSIIQSMGYRNRAKRLQSEPNPSLQETSLSVQSLNIGT 161
QY 96 QODGLEKHLPEVRIP-HMERGESAVMLEPRIKIAILGLSS--IGTPEGITAEVLV 152
DB 162 VRTLRTKRIOPQKTSYVIELGSDS---EDTVNKATYCSVGDQLQITPGTRDEI-- 216
QY 153 VTSFDELORASE 165
DB 217 --SLDSAKKACE 227

RESULT 9

US-09-734-672-6
; Sequence 6, Application US/09734672
; Publication No. US20020183268A1
GENERAL INFORMATION:
APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvares, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Schelter, Denise B.
Zeng, Bin
TITLE OF INVENTION: Coding Sequences of the Human
BRCA1 Gene
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan Lewis & Bockius LLP
STREET: 1111 Pennsylvania Ave., N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/734,672
FILING DATE: 03-Dec-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/966,436
FILING DATE: 07-No. US20020183268A1-97
APPLICATION NUMBER: US 08/598,591
FILING DATE: 12-Feb-96
ATTORNEY/AGENT INFORMATION:
NAME: Michael S. Tuscan
REGISTRATION NUMBER: 43,210
REFERENCE/DOCKET NUMBER: 44921-5055-02-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-739-3000
TELEFAX: 202-739-3001
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 amino acids
TYPE: amino acid
STRANDEDNESS: NO. US20020183268A1 Relevant
TOPOLOGY: NO. US20020183268A1 Relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: BRCA1
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 17
MAP POSITION: 17q21
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-734-672-6
Query Match 8.9%; Score 78.5; DB 10; Length 1863;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;
QY 5 IFAPFGVHLISLCSGKA---ICNGISKRT-----FEIKERIAS--CG---DVAKAI 50
DB 42 IFCKFCMLKLNQKGBSQPLCKNDITKRSLOESTRFSQVLEBLKTIICAFOLDTGLEV 101
QY 51 INLAVYGAQRNSYERL---ALLVTVGPRLSGSKNL-----EKAQIMYQNL-- 95
DB 102 ANSNVPAKKNNSPRLKDEVSIIQSMGYRRAKRLQSEPDNSLOETSLSVGLSNIGT 161
QY 96 QODGLEKHLKPRVRIIP-HWEGEBSAVMLERIRIKIALIGLGS--IGTPPEGITAEVLV 152
DB 162 VRLTRTKRRIOPKTSVYIEIGSDSS---EDTVNKATYCSVGDDDELLQITPQGRDEI-- 216

QY 153 VTSFDELORRASE 165
DB 217 --SIDSAXKXACE 227
RESULT 10
US-09-982-828-2
; Sequence 2, Application US/09982828
; Publication No. US20030022184A1
GENERAL INFORMATION:
APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvares, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Thurber, Denise
Zeng, Bin
TITLE OF INVENTION: Coding Sequences of the Human
BRCA1 Gene
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan Lewis & Bockius LLP
STREET: 1111 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,828
FILING DATE: 22-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/074,453
FILING DATE: 1998-05-06
APPLICATION NUMBER: US 08/798,691
FILING DATE: 1997-02-12
APPLICATION NUMBER: US 08/598,591
FILING DATE: 1996-02-12
ATTORNEY/AGENT INFORMATION:
NAME: Michael S. Tuscan
REGISTRATION NUMBER: 43,210
REFERENCE/DOCKET NUMBER: 44921-5053-01-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-739-3000
TELEFAX: 202-739-3001
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: BRCA1 (cm11)
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 17
MAP POSITION: 17q21
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-982-828-2
Query Match 8.9%; Score 78.5; DB 11; Length 1863;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;
QY 5 IFAPFGVHLISLCSGKA---ICNGISKRT-----FEIKERIAS--CG---DVAKAI 50
DB 42 IFCKFCMLKLNQKGBSQPLCKNDITKRSLOESTRFSQVLEBLKTIICAFOLDTGLEV 101

QY 51 INLA VYGAQNRSYERL---ALVDYVGPRLSGSKNL-----EKAIQIMYQNL-- 95
DB 102 ANSYNPAKKNNSPEHLKDEVSIIOGMYRNARAKLLQSEPNPBLQSTSLSVQLSNIGT 161
QY 96 QODGLEKVLHLEVPRI-P-HMERGESAVMLBPRIKHIAIIGLSS--IGTPPEGITAEVLV 152
DB 162 VRTLRTRKRIOPQKTSYVIELGSDS---EDTVNKATYCSVGDOELLQITPQGTREDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKAACE 227

RESULT 11
US-09-982-828-4
Sequence 4, Application US/09982828
Publication No. US20030022184A1
GENERAL INFORMATION:
APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvarez, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Thurber, Denise
Zeng, Bin
TITLE OF INVENTION: Coding Sequences of the Human
BRCA1 Gene
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan Lewis & Bockius LLP
STREET: 1111 Pennsylvania Avenue N. W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,828
FILING DATE: 22-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/074,453
FILING DATE: 1998-05-06
APPLICATION NUMBER: US 08/798,691
FILING DATE: 1997-02-12
APPLICATION NUMBER: US 08/598,591
FILING DATE: 1996-02-12
ATTORNEY/AGENT INFORMATION:
NAME: Michael S. Tuscan
REGISTRATION NUMBER: 43,210
REFERENCE/DOCKET NUMBER: 44921-5053-01-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-739-3000
TELEFAX: 202-739-3001
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: Protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: BRCA1 (om12)
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 17
MAP POSITION: 17q21
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-982-828-4
Query Match 8.9%; Score 78.5; DB 11; Length 1663;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;
QY 51 IFAFGVHLLSLCSGKA---ICKNGISKRT-----FEEIKERIAS--CG--DVAKAI 50
DB 42 IFCKFCMLKILNKGKGSQCPLCNQDITKRSIQESTRFSQVLEBLKICAFQDGTGLEV 101
QY 51 INLA VYGAQNRSYERL---ALVDYVGPRLSGSKNL-----EKAIQIMYQNL-- 95
DB 102 ANSYNPAKKNNSPEHLKDEVSIIOGMYRNARAKLLQSEPNPBLQSTSLSVQLSNIGT 161
QY 96 QODGLEKVLHLEVPRI-P-HMERGESAVMLBPRIKHIAIIGLSS--IGTPPEGITAEVLV 152
DB 162 VRTLRTRKRIOPQKTSYVIELGSDS---EDTVNKATYCSVGDOELLQITPQGTREDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKAACE 227

RESULT 12
US-09-982-828-6
Sequence 6, Application US/09982828
Publication No. US20030022184A1
GENERAL INFORMATION:
APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvarez, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Thurber, Denise
Zeng, Bin
TITLE OF INVENTION: Coding Sequences of the Human
BRCA1 Gene
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan Lewis & Bockius LLP
STREET: 1111 Pennsylvania Avenue N. W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,828
FILING DATE: 22-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/074,453
FILING DATE: 1998-05-06
APPLICATION NUMBER: US 08/798,691
FILING DATE: 1997-02-12
APPLICATION NUMBER: US 08/598,591
FILING DATE: 1996-02-12
ATTORNEY/AGENT INFORMATION:
NAME: Michael S. Tuscan
REGISTRATION NUMBER: 43,210
REFERENCE/DOCKET NUMBER: 44921-5053-01-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-739-3000
TELEFAX: 202-739-3001
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant

```

/
/      TOPOLOGY: not relevant
/      MOLECULE TYPE: protein
/      ORIGINAL SOURCE:
/      ORGANISM: Homo sapiens
/      STRAIN: BRCAL (om13)
/      POSITION IN GENOME:
/      CHROMOSOME/SEGMENT: 17
/      MAP POSITION: 17q21
/      SEQUENCE DESCRIPTION: SEQ ID NO: 6:
/
US-09-982-828-6

Query Match      8.9%; Score 78.5; DB 11; Length 1863;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGGVHLLSLCSGKA---ICNGISKRT-----FEEIKELIAS--CG---DVAKAI 50
DB 42 IFCKFCMLKILNKKQKSPQCPCKNDITKRSLOSTFSSQLVEBLKIIICAFOLDTGLE 101
QY 51 INLAVGAQNRSEYERL---ALVDTVGPRLSGSKNL-----EKAQIMYQNL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNRAKRLQSEPNLSLOETSLSVQLSNLGT 161
QY 96 QODGLEKVLHPVRIP-HMERGESAVMLEPRIHKIALIQLGSS--IGTPEGITAEVLV 152
DB 162 VRTLRTKQRIQPKQTSVYIELGSDS---EDTVNKATYCSVGQDELQITPGTRDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKACE 227

RESULT 13
US-10-022-819-2
/ Sequence 2, Application US/10022819
/ Publication No. US20030027166A1
/ GENERAL INFORMATION:
/ APPLICANT: ALLEN, Antonette C. P.
/ OLESEN, Sheri J.
/ LAWRENCE, Tammy
/ ANGELLY, Tracy S.
/ RABIN, Mark B.
/ TITLE OF INVENTION: CODING SEQUENCE HAPLOTYPE OF THE HUMAN
/ BRCAL GENE
/ NUMBER OF SEQUENCES: 67
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: Morgan Lewis & Bockius LLP
/ STREET: 1111 Pennsylvania Avenue
/ CITY: Washington DC
/ STATE: District of Columbia
/ COUNTRY: USA
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/022,819
/ FILING DATE: 22-Apr-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/074,452
/ FILING DATE: 1998-05-06
/ ATTORNEY/AGENT INFORMATION:
/ NAME: <Unknown>
/ REGISTRATION NUMBER: <Unknown>
/ REFERENCE/DOCKET NUMBER: 044921-5049-01-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-739-3000
/ TELEFAX: 202-739-3001
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:

```

```

/
/      LENGTH: 1863 amino acids
/      TYPE: amino acid
/      STRANDEDNESS: single
/      TOPOLOGY: linear
/      MOLECULE TYPE: protein
/      HYPOTHETICAL: NO
/      ANTI-SENSE: NO
/      FRAGMENT TYPE: N-terminal
/      ORIGINAL SOURCE:
/      ORGANISM: HOMO SAPIENS
/      STRAIN: BRCAL
/      HAPLOTYPE: OM14
/      POSITION IN GENOME:
/      CHROMOSOME/SEGMENT: 17
/      MAP POSITION: 17q21
/      SEQUENCE DESCRIPTION: SEQ ID NO: 2:
/
US-10-022-819-2

Query Match      8.9%; Score 78.5; DB 15; Length 1863;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGGVHLLSLCSGKA---ICNGISKRT-----FEEIKELIAS--CG---DVAKAI 50
DB 42 IFCKFCMLKILNKKQKSPQCPCKNDITKRSLOSTFSSQLVEBLKIIICAFOLDTGLE 101
QY 51 INLAVGAQNRSEYERL---ALVDTVGPRLSGSKNL-----EKAQIMYQNL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNRAKRLQSEPNLSLOETSLSVQLSNLGT 161
QY 96 QODGLEKVLHPVRIP-HMERGESAVMLEPRIHKIALIQLGSS--IGTPEGITAEVLV 152
DB 162 VRTLRTKQRIQPKQTSVYIELGSDS---EDTVNKATYCSVGQDELQITPGTRDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKACE 227

RESULT 14
US-09-925-299-1011
/ Sequence 1011, Application US/09925299
/ Patent No. US20020055627A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA102
/ CURRENT APPLICATION NUMBER: US/09/925,299
/ CURRENT FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05883
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 1596
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 1011
/ LENGTH: 384
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-925-299-1011

Query Match      8.4%; Score 74; DB 9; Length 384;
Best Local Similarity 26.0%; Pred. No. 12;
Matches 34; Conservative 24; Mismatches 43; Indels 30; Gaps 8;

QY 23 ICNGISKRTFEEIKELIASCG--DVAKAIINLAV---YGR--AQNRSEYER-LALLVDVT 73
DB 82 LCGRAL-----EIREKVLGTNNHPDVAKQNLNLLALCQNGKKEAVERRYQRLALIVEGQ 135
QY 74 VG---PRLSGSKN-----LEKAIQIMYQNLQODGLEKVLHPV-----RIPMERGES 118
DB 136 LGPDNPVAVARTKNNLASCYLKQGYAEAEFTLYKEILTRAHVQOERGSVDDHKPTWMAEE 195

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:50:05 ; Search time 41 Seconds
(without alignments)
38.714 Million cell updates/sec

Title: US-09-745-763-36_COPY_231_240
Perfect score: 10
Sequence: 1 TACITVEDAE 10

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	100.0	178	24	ABR48107	Human secreted pro
2	100.0	178	24	ABR00275	Human gene 142 enc
3	100.0	472	18	AAW33604	Human secreted pro
4	100.0	472	19	AAW85456	Secreted protein e
5	100.0	472	21	AAH10229	Human fetal kidney
6	100.0	472	21	AAH7254	Human signal pepti
7	100.0	472	21	AAH58879	Human peptidase NA
8	100.0	472	22	AAW93559	Human polypeptide,
9	100.0	472	22	AAW60558	Human gene 18 enco

10	100.0	472	23	ABP61800	Human polypeptide
11	100.0	472	23	ABG33880	Human secreted pro
12	100.0	472	24	ABR47892	Human secreted pro
13	100.0	472	24	ABR48106	Human secreted pro
14	100.0	472	24	ABR00152	Human gene 142 enc
15	100.0	472	24	ABR00275	Human gene 142 enc
16	100.0	473	21	AAH7081	Human secreted pro
17	100.0	474	22	ABG09405	Novel human diago
18	100.0	481	22	AAH25762	Human protein sequ
19	60.0	288	22	AAH81198	Human RNA-binding
20	60.0	321	22	ABG22827	Novel human diago
21	60.0	323	22	AAH93971	Human stomach cano
22	60.0	323	22	AAH93767	Human protein sequ
23	60.0	325	21	AAH95335	Human quaking spli
24	60.0	328	23	ABP66344	Bifidobacterium 10
25	60.0	341	21	AAH95336	Human quaking spli
26	60.0	344	21	AAH06629	Arabidopsis thalia
27	60.0	344	21	AAH48593	Arabidopsis thalia
28	60.0	377	22	AAH62340	Arabidopsis thalia
29	60.0	417	21	AAH07093	Propionibacterium
30	60.0	417	21	AAH49104	Arabidopsis thalia
31	60.0	424	21	AAH07092	Arabidopsis thalia
32	60.0	424	21	AAH49103	Arabidopsis thalia
33	60.0	425	21	AAH77270	Streptomyces pence
34	60.0	425	21	AAH78822	Amino acid sequenc
35	60.0	432	22	AAH36107	Klebsiella pneumon
36	60.0	433	21	AAH19094	Protein encoded by
37	60.0	448	21	AAH07091	Arabidopsis thalia
38	60.0	448	21	AAH49102	Arabidopsis thalia
39	60.0	532	20	AAH16036	Neurospora crassa
40	60.0	610	22	AAH34063	Staphylococcus aur
41	60.0	618	22	AAH36974	Staphylococcus aur
42	60.0	618	22	AAH37377	Staphylococcus aur
43	60.0	671	23	AAH49703	Listeria monocytog
44	60.0	1068	20	AAH30048	Pancreatic eukaryo
45	60.0	1108	20	AAH30046	Pancreatic eukaryo

ALIGNMENTS

RESULT 1	
ID	ABR48107
ABR48107	standard; Protein; 178 AA.
AC	ABR48107;
DT	12-JUN-2003 (first entry)
XX	
DE	Human secreted protein, SEQ ID 998.
XX	
KW	Cardiant; antiarrhythmic; antiarteriosclerotic; vasotrophic; cyostatic;
KW	vulnerary; antiinflammatory; nocotropic; neuroprotective;
KW	antiparkinsonian; gene therapy; human; cardiovascular disorder.
XX	
OS	Homo sapiens.
XX	
FN	W0200295010-A2.
XX	
PD	28-NOV-2002.
XX	
PF	19-MAR-2002; 2002W0-US09785.
XX	
PR	21-MAR-2001; 2001US-277340P.
PR	19-JUL-2001; 2001US-306171P.
XX	13-NOV-2001; 2001US-331287P.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Ruben SM;
XX	
DR	WPI; 2003-129429/12.
XX	

PT Novel human secreted proteins, useful for detecting, preventing,
PT diagnosing, prognosticating, treating and/or ameliorating
PT cardiovascular disorders such as arrhythmia -
XX
XX Claim 13; SEQ ID 998, 1881pp; English.
XX
XX The present invention relates to novel human secreted proteins
CC (ABR47733-ABR48145) and their coding sequences (ACC50344-ACC50856). The
CC proteins and their coding sequences are useful for the preparation of a
CC diagnostic or pharmaceutical composition for diagnosing or treating a
CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
CC coronary arteriosclerosis and myocardial ischemia), neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, renal disorders,
CC proliferative disorders and/or cancerous diseases and conditions, for
CC wound healing and epithelial cell proliferation, to treat inflammation or
CC infection, for treating thrombosis and arteriosclerosis, for treating or
CC preventing neural damage which occurs in neuronal disorders or
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
CC disease, to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts, to prevent skin aging or hair loss, to
CC stimulate growth and differentiation of haematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism.
CC Note: The sequence data for this patent was published in electronic
CC format and is available from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 178 AA;

Query Match 100.0%; Score 10; DB 24; Length 178;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TACTIVEDAE 10
Db 21 TACTIVEDAE 30

RESULT 2
ABR00275
ID ABR00275 standard; Protein; 178 AA.
XX
AC ABR00275;
XX
DT 03-APR-2003 (first entry)
XX
DE Human gene 142 encoded secreted protein HRACJ35, SEQ ID NO:564.
XX
XX Human, secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic;
KW antiinflammatory; immunosuppressive; vulnerrary; gene therapy.
XX
OS Homo sapiens.
XX
XX WO200276488-A1.
XX
XX PD 03-OCT-2002.
XX
XX 19-MAR-2002; 2002WO-US08276.
XX
XX PF 21-MAR-2001; 2001US-277340P.
XX
XX PR 19-JUL-2001; 2001US-306171P.
XX
XX PR 13-NOV-2001; 2001US-33187P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX

PI Rosen CA, Ruben SM;
XX
XX WPI; 2003-029900/02.
DR N-PSDB; AB271454.
XX
XX New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers -
XX
XX Claim 13; Page 1110; 1216pp; English.
XX
XX AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
CC AB271479-AB271540 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening, and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing, treating, ameliorating or preventing
CC digestive disorders. Such conditions include disorders of the mouth,
CC oesophagus, stomach, small intestine, large intestine, liver, biliary
CC tract and pancreas, and include cancers of these organs and tissues. The
CC secreted proteins and their nucleic acids may also be used in the
CC treatment of immune disorders, inflammation, infection,
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
CC of the invention may be used for chromosome identification, chromosome
CC mapping, in gene therapy, for identifying individuals from minute
CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein of the
CC invention.
XX
SQ Sequence 178 AA;

Query Match 100.0%; Score 10; DB 24; Length 178;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TACTIVEDAE 10
Db 21 TACTIVEDAE 30

RESULT 3
AAW33604
ID AAW33604 standard; Protein; 472 AA.
XX
XX AAW33604;
XX
DT 21-MAY-1998 (first entry)
XX
DE Human secreted protein AM282 full-length sequence.
XX
XX Secreted protein; AM282; cytokine; human.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..24
FT /label= Sig_peptide
FT 25..472
FT Protein /label= Mat_protein
XX
XX W09739030-A2.
XX
XX PD 23-OCT-1997.
XX
XX PF 16-APR-1997; 97WO-US06475.
XX
XX PR 13-JAN-1997; 97US-0783520.
XX
XX PR 18-APR-1996; 96US-0634325.
XX
XX (GEMY) GENETICS INST INC.
XX

XX	Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;
PI	Spaulding V;
XX	
XX	WPI; 1997-526400/48.
DR	N-PSDB; AAV02296.
XX	
PT	New isolated secretory proteins AM340, AM282 and AK583 - possibly
PT	have cytokine, cell proliferation/differentiation regulating,
PT	immunomodulating activities, etc.
XX	
PS	Claim 16; Page 45-47; 59pp; English.
XX	
CC	This human secreted protein, designated AM282, is encoded by a
CC	full-length cDNA clone (see AAV02296), deposited in ATCC 98026, that
CC	was identified from a database search using an isolated partial
CC	AM282 clone (see AAV97398). AM282 protein can be used in a claimed
CC	method for preventing, treating or ameliorating a medical
CC	condition. It may exhibit cytokine, cell proliferation (either
CC	inducing or inhibiting) or cell differentiation (either inducing or
CC	inhibiting) activity or may induce production of other cytokines in
CC	certain cell populations. It may also exhibit e.g. immune
CC	stimulating or suppressing activity, haematopoiesis regulating
CC	activity, tissue growth activity, activin/inhibin activity,
CC	chemotactic or chemokinetic activity, haemostatic or thrombolytic
CC	activity, receptor/ligand activity, anti-inflammatory activity, of
CC	tumour inhibition activity, or other activities. No evidence of
CC	any of these activities is given in the specification.
XX	
SO	Sequence 472 AA;
	Query Match 100.0%; Score 10; DB 18; Length 472;
	Best Local Similarity 100.0%; Pred. No. 0.0051;
	Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 TACTTVEADAE 10
Db	231 TACTTVEADAE 240
RESULT 4	
ID	AAW85456 standard; Protein; 472 AA.
AAW85456	
AC	AAW85456;
XX	
DT	25-FEB-1999 (first entry)
XX	
DB	Secreted protein encoded by clone bu45_2.
XX	
KM	Secreted protein; nutritional activity; immune stimulating; vaccine;
KM	suppressing activity; haematopoiesis regulating activity;
KM	tissue growth activity; activin; inhibin activity; chemotaxis;
KM	chemokinetic activity; haemostasis; thrombolytic activity; receptor;
KM	ligand; anti-inflammatory; cadherin; tumour invasion suppressor;
KM	tumour inhibition; gene therapy.
XX	
OS	Homo sapiens.
XX	
PN	WO9842739-A2.
XX	
PD	01-OCT-1998.
XX	
PF	20-MAR-1998; 98WO-US05653.
XX	
PR	19-MAR-1998; 98US-0044466.
XX	
PR	21-MAR-1997; 97US-0822167.
XX	
PA	(GENY) GENETICS INST INC.
XX	
PI	Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI	Racie LA, Spaulding V, Treacy M;
XX	

DR WP1: 1998-609890/51.
DR N-PSDB; AAV62779.
PT New polynucleotides encoding secreted human proteins - derived from
PT human foetal brain, adult brain, foetal kidney, placenta or adult
PT pineal gland cDNA libraries.
XX
XX
PS Claim 14; Page 70-72; 113pp; English.
XX
XX The present sequence represents a secreted protein. The polynucleotide
CC and secreted protein are predicted to have biological activities which
CC would make them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals, although no supporting data is given.
CC Suggested activities include nutritional activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity (no data is
CC given in the specification to support these activities). The
CC polynucleotide is also stated to be useful for gene therapy.
XX
XX
SQ Sequence 472 AA;

Query Match 100.0%; Score 10; DB 19; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TACTTVEDAE 10
|||||||
Db 231 TACTTVEDAE 240

RESULT 5
AAB10229
ID AAB10229 standard; Protein: 472 AA.
XX
XX AAB10229;
AC
XX
XX 16-NOV-2000 (first entry)
DT
XX
XX Human fetal kidney protein fragment AM282_11.
DE
XX Secreted protein; cytosolic; immunostimulatory; antimicrobial;
KW antiviral; immunosuppressive; antiinflammatory; vulnerrary; cytokine;
KW cell proliferation; differentiation; regulator; treatment; tumor;
KW autoimmune disease; inflammatory disorder; wound; microbial infection;
KW viral disease; graft versus host reaction suppression.
XX
XX Homo sapiens.
OS
XX WO200037630-A1.
PN
XX 29-JUN-2000.
PD
XX 22-DEC-1999; 99WO-US31005.
PF
XX 23-DEC-1998; 98US-0220876.
PR
XX (GENY) GENETICS INST INC.
PA
XX Jacobs K, McCoy JM, Lavalie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR;
XX
XX WP1: 2000-442661/38.
DR N-PSDB; AAA40493.
XX
XX Secreted human proteins AS296-1i and AS34-1i, useful for treating
PT tumors, autoimmune diseases, inflammatory disorders, wounds, microbial
PT infections and viral diseases -
XX
XX Disclosure; Page 199-200; 293pp; English.
XX

XX (JANC) JANSSEN PHARM NV.
 XX Pangalos M, Neefs JEFM, Peeters DCG;
 XX WPI, 2000-182424/16.
 XX N-PSDB; AAZ58313.
 XX New human N-acetylated alpha-linked acidic dipeptidases for treating
 PT neural disorders e.g. Alzheimer's disease, schizophrenia and
 PT Parkinson's disease -
 XX Claim 15; Fig 5; 95bp; English.
 XX The present sequence is that predicted for human N-acetylated
 CC alpha-linked acidic dipeptidase IV (NALAD-ase IV) on the basis of
 CC isolated cDNA (see AAZ58313). NALAD-ase IV is predicted to be a
 CC type II integral membrane protein of mol.wt. 51.9 kDa and pI 5.99.
 CC Expression was low in all tissues examined by RT-PCR. The
 CC invention provides human NALAD-ase I, II and IV cDNAs and encoded
 CC polypeptides, as well as vectors, host cells, transgenic organisms,
 CC antisense nucleic acids, agonists and antagonists. These are useful
 CC for treating neural disorders such as Alzheimer's disease,
 CC schizophrenia, ALS, Parkinson's disease, peripheral neuropathy,
 CC Huntington's disease, acute brain injury, multiple sclerosis,
 CC exposure to neurotoxins, peripheral nerve trauma, ischaemia or
 CC dementia (claimed).
 XX Sequence 472 AA;
 SQ
 Query Match 100.0%; Score 10; DB 21; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TACTTVEDAE 10
 DB 231 TACTTVEDAE 240
 RESULT 8
 AAM93559 standard; Protein; 472 AA.
 XX AAM93559;
 AC AAM93559;
 XX 06-NOV-2001 (first entry)
 DT Human polypeptide, SEQ ID NO: 3329.
 DE Human polypeptide, CDNA synthesis; oligo-capping.
 XX Human, full length CDNA; CDNA synthesis; oligo-capping.
 KM Homo sapiens.
 XX EP1130094-A2.
 XX EP1130094-A2.
 PD 05-SEP-2001.
 XX 07-JUL-2000; 2000EP-0114089.
 PF 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX (HELI-) HELIX RES INST.
 PA Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 DR N-PSDB; AAK94491.
 XX 830 Primers useful for synthesizing full length CDNA clones and their
 PT in genetic manipulation -

XX Claim 8; SEQ ID NO 3329; 1380bp + sequence listing; English.
 XX The invention relates to primers for synthesizing full length CDNA
 CC clones. 830 CDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5' - and 3' -ends of the CDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC CDNA are useful for clarifying the function of the protein encoded by
 CC the CDNA. The full length clones were obtained by construction of full
 CC length enriched CDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length CDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human CDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX Sequence 472 AA;
 SQ
 Query Match 100.0%; Score 10; DB 22; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TACTTVEDAE 10
 DB 231 TACTTVEDAE 240
 RESULT 9
 AAE06058 standard; Protein; 472 AA.
 XX AAE06058;
 AC AAE06058;
 XX 24-SEP-2001 (first entry)
 DT Human gene 18 encoded secreted protein HRA0335, SEQ ID NO:120.
 DE Human gene 18 encoded secreted protein HRA0335, SEQ ID NO:120.
 XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 KM foetal abnormality; developmental abnormality; hematopoietic disorder;
 KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KM Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KM psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KM inflammation; neurological disorder; Alzheimer's disease; food additive;
 KM angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 KM pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KM cell culture; chemotaxis; vulnery; binding partner identification;
 KM gene therapy.
 XX Homo sapiens.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal_peptide
 FT 25..472
 FT Protein /label= Mature_human_secreted_protein
 XX W0200151504-A1.
 XX 19-JUL-2001.
 PD 12-JAN-2001; 2001WO-US00911.
 PF 13-JAN-2000; 2000US-0482273.
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Endress GA, Muscenski M, Eimer R;
 XX WPI; 2001-425865/45.
 DR N-PSDB; AAD11647.
 XX

PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 XX
 PS Claim 11; Page 750-752; 864pp; English.
 XX
 CC AAD1630-AAD11721 represent cDNAs corresponding to 71 human secreted
 CC protein genes, and AAB06041-AAB06132 represent the proteins they encode.
 CC AAB06133-AAB06205 represent human secreted protein fragments.
 CC The secreted proteins and their genes are useful for preventing, treating
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 CC Pathological conditions can be diagnosed by determining the amount of the
 CC new protein in a sample or by determining the presence of mutations in
 CC the new genes. Specific uses are described for each of the 71 genes,
 CC based on the tissues in which they are most highly expressed, and include
 CC developing products for the diagnosis or treatment of proliferative
 CC disorders, cancer, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
 CC pregnancy-related disorders, endocrine disorders, and infections. The
 CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, for supporting cell culture of primary tissues,
 CC to regenerate tissues, to identify their cognate ligands or binding
 CC partners, and in chemotaxis, and can be used as a food additive or
 CC preservative to modify storage properties. Antibodies specific for a
 CC protein of the invention can be used in alleviating symptoms associated
 CC with the disorders mentioned above, and in diagnostic immunoassays (e.g.,
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA)).
 CC The present sequence represents a human secreted protein of
 CC the invention.
 XX
 SQ Sequence 472 AA;
 Query Match 100.0%; Score 10; DB 22; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TACTIVEDAE 10
 |||||
 Db 231 TACTIVEDAE 240
 RESULT 10
 ABB61800
 ID ABB61800 standard; Protein; 472 AA.
 XX
 AC ABB61800;
 XX
 XX 04-OCT-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 154.
 XX
 KM Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic;
 KM antinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
 KM neuroprotective; nociceptive; osteopathic; haemostatic; vasotrophic;
 KM antifungal; fungicide; antidiabetic; antiallergic;
 KM immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KM cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KM stem cell; growth factor; nervous system disease; neuropathy;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KM osteoporosis; severe combined immunodeficiency; SCID; infection;
 KM multiple sclerosis; rheumatoid arthritis; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN US2002065394-A1.
 XX
 PD 30-MAY-2002.
 XX

PF 22-DEC-2000; 2000US-0745763.
 XX
 PR 18-MAR-1998; 98US-0040963.
 XX
 PA (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREB/) TREACY M.
 PA (SPAU/) SPAULDING V.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 XX
 DR WPI; 2002-582343/62.
 DR N-PSDB; ABQ92016.
 XX
 PT Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis
 PT
 PS Claim 51; Page 114-115; 284pp; English.
 XX
 CC The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (cDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (I) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polypeptide of the invention.
 XX
 SQ Sequence 472 AA;
 Query Match 100.0%; Score 10; DB 23; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TACTIVEDAE 10
 |||||
 Db 231 TACTIVEDAE 240
 RESULT 11
 ABB33880
 ID ABB33880 standard; Protein; 472 AA.
 XX

AC ABG33880;
 XX 15-JUL-2002 (first entry)
 XX
 DE Human secreted protein encoded by gene 18 #1.
 XX
 KW Human; secreted protein; gene therapy; immunosuppressive;
 KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiac;
 KW vasotropic; cerebroprotective; nocotropic; neuroprotective; antibacterial;
 KW virocidic; fungicide; ophthalmological; autoimmune disease; neoplasm;
 KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
 KW ocular disorder; corneal infection; wound healing; skin aging;
 KW epithelial cell proliferation; food additive.
 XX
 OS Homo sapiens.
 PN WO200226931-A2.
 PD 04-APR-2002.
 PF 24-SEP-2001; 2001WO-US29871.
 PR 25-SEP-2000; 2000US-234925P.
 PR 12-JAN-2001; 2001WO-US00911.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y,
 PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR,
 PI Endress GA, Mucenski M, Ebner R;
 XX
 DR WPI: 2002-362489/39.
 DR N-PSDB; ABK69743.
 XX
 PT Novel 71 isolated secreted polypeptides and polynucleotides encoding
 PT the polypeptides, useful for treating Huntington's disease, sepsis,
 PT meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
 PT asthma
 PT
 PS Claim 11; Page 1231-1232; 1478pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (or its
 CC fragment, homologue complement or allelic variant) encoding a human
 CC secreted protein (and its fragment, domain, epitope, variant, secreted
 CC form and species variant). Also included are a recombinant vector
 CC comprising the nucleic acid, a recombinant host cell comprising the
 CC vector, an antibody against the secreted protein, a recombinant host cell
 CC that expresses the secreted protein and a method of identifying a binding
 CC partner of the secreted protein. The nucleic acid and protein are used to
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
 CC human, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
 CC for example autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
 CC ocular disorders e.g. corneal infection. Many other diseases and
 CC disorders are listed in the specification. The polypeptides can also be
 CC used to aid wound healing an epithelial cell proliferation, to prevent
 CC skin aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities. The present
 CC sequence represents a novel human secreted protein of the invention.
 XX
 SQ Sequence 472 AA;

Query Match 100.0%; Score 10; DB 23; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACTIVEDAE 10
 |||||
 Db 231 TACTIVEDAE 240

RESULT 12
 ABR47892
 ID ABR47892 standard; Protein; 472 AA.
 XX
 AC ABR47892;
 XX
 DT 12-JUN-2003 (first entry)
 XX
 DE Human secreted protein, SEQ ID 783.
 XX
 KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
 KW vulnerary; antiinflammatory; nocotropic; neuroprotective;
 KW antiparkinsonian; gene therapy; human; cardiovascular disorder.
 XX
 OS Homo sapiens.
 PN WO200295010-A2.
 PD 28-NOV-2002.
 PF 19-MAR-2002; 2002WO-US09785.
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI: 2003-129429/12.
 DR
 XX
 PT Novel human secreted proteins, useful for detecting, preventing,
 PT diagnosing, prognosticating, treating and/or ameliorating
 PT cardiovascular disorders such as arrhythmia
 PT
 PS Claim 13; SEQ ID 783; 1881pp; English.
 XX
 CC The present invention relates to novel human secreted proteins
 CC (ABR47892-ABR48145) and their coding sequences (ACC50344-ACC50856). The
 CC proteins and their coding sequences are useful for the preparation of a
 CC diagnostic or pharmaceutical composition for diagnosing or treating a
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
 CC coronary arteriosclerosis and myocardial ischaemia), neural disorders,
 CC immune system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of
 CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism.
 CC Note: The sequence data for this patent was published in electronic
 CC format and is available from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 472 AA;

Query Match 100.0%; Score 10; DB 24; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TACTIVEDAE 10
 |||||
 Db 231 TACTIVEDAE 240

RESULT 13

ABR48106
 ID ABR48106 standard; Protein; 472 AA.

AC ABR48106;

DT 12-JUN-2003 (first entry)

DE Human secreted protein, SEQ ID 997.

XX Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;
 KW vulnerable; antiinflammatory; nootropic; neuroprotective;
 KW antiparkinsonian; gene therapy; human; cardiovascular disorder.

XX Homo sapiens.

XX WO200295010-A2.

PD 28-NOV-2002.

PE 19-MAR-2002; 2002WO-US09785.

XX 21-MAR-2001; 2001US-277340P.

PR 19-JUL-2001; 2001US-306171P.

PR 13-NOV-2001; 2001US-331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI; 2003-129429/12.

XX Novel human secreted proteins, useful for detecting, preventing,
 PT diagnosing, prognosticating, treating and/or ameliorating
 PT cardiovascular disorders such as arrhythmia -

PS Claim 13; SEQ ID 997; 1881pp; English.

XX The present invention relates to novel human secreted proteins
 CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The
 CC proteins and their coding sequences are useful for the preparation of a
 CC diagnostic or pharmaceutical composition for diagnosing or treating a
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
 CC coronary arteriosclerosis and myocardial ischemia), neural disorders,
 CC immune system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of haematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of
 CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism.
 CC Note: The sequence data for this patent was published in electronic
 CC format and is available from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 472 AA;

Query Match 100.0%; Score 10; DB 24; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.0051;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 TACTIVEDAE 10
 |||||
 Db 231 TACTIVEDAE 240

RESULT 14

ABR00152
 ID ABR00152 standard; Protein; 472 AA.

AC ABR00152;

DT 03-APR-2003 (first entry)

DE Human gene 142 encoded secreted protein HRA035, SEQ ID NO:441.

XX Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
 KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
 KW antiinflammatory; immunosuppressive; vulnerable; chromosome 8q22.2.

XX Homo sapiens.

XX WO200276488-A1.

XX 03-OCT-2002.

PE 19-MAR-2002; 2002WO-US08276.

XX 21-MAR-2001; 2001US-277340P.

PR 19-JUL-2001; 2001US-306171P.

PR 13-NOV-2001; 2001US-331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI; 2003-029900/02.

DR N-PSDB; ABR71331.

XX New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 PT e.g. gastrointestinal diseases and disorders, or cancers -
 XX Claim 13; Page 1029-1030; 1216pp; English.

XX ABR71190-ABR71478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABR00011-ABP00299 represent the proteins they encode.
 CC ABR71479-ABR71540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein of the
 CC invention.

XX Sequence 472 AA;

Query Match 100.0%; Score 10; DB 24; Length 472;

Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TACTIVEDAE 10
 |||||
 DB 231 TACTIVEDAE 240

Query Match 100.0%; Score 10; DB 24; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TACTIVEDAE 10
 |||||
 DB 231 TACTIVEDAE 240

Search completed: December 22, 2003, 11:56:04
 Job time : 42 secs

RESULT 15
 ABR00274
 ID ABR00274 standard; Protein: 472 AA.
 AC ABR00274;
 DT 03-APR-2003 (first entry).
 XX
 DE Human gene 142 encoded secreted protein HRA035, SEQ ID NO:563.
 XX
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
 KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
 KW antiinflammatory; immunosuppressive; vulnerary; Chromosome 8q22.2.
 XX
 OS Homo sapiens.
 XX
 PN WO200276488-A1.
 XX
 PD 03-OCT-2002.
 XX
 PF 19-MAR-2002; 2002WO-US08276.
 XX
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-029900/02.
 DR N-PSDB; ABZ71453.
 XX
 PT New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating,
 PT e.g. gastrointestinal diseases and disorders, or cancers -
 XX
 PS Claim 13; Page 1108-1109; 1216pp; English.
 XX
 CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein of the
 CC invention.
 XX
 SQ Sequence 472 AA;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:54:11 ; Search time 21 Seconds
(without alignments)
45.795 Million cell updates/sec

Title: US-09-745-763-36_COPY_231_240
Perfect score: 10
Sequence: 1 TACTIVEDAE 10

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : PIR_76:*

1: p1r1:*\n2: p1r2:*\n3: p1r3:*\n4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.0	157	2	T05516	hypothetical prote
2	60.0	171	2	E82459	probable acetyltra
3	60.0	187	2	E87299	hypothetical prote
4	60.0	197	2	F82451	probable acetyltra
5	60.0	268	2	J01473	pancreatic elastase
6	60.0	333	2	F83331	probable transcrip
7	60.0	334	2	G71684	virB11 protein (vi
8	60.0	334	2	F97748	virB11 protein (im
9	60.0	344	2	T02714	hypothetical prote
10	60.0	348	2	E83920	hypothetical prote
11	60.0	349	2	T24744	hypothetical prote
12	60.0	362	2	A47211	protein kinase BRK
13	60.0	369	2	A60041	G2+/calmodulin-de
14	60.0	371	2	S51320	mitogen-activated
15	60.0	387	2	S48123	mitogen-activated
16	60.0	393	2	S51321	mitogen-activated
17	60.0	433	2	S37790	probable serine/th
18	60.0	455	2	F82345	conserved hypotbet
19	60.0	465	2	J04793	dihydroliponamide d
20	60.0	532	2	T46649	1,3-beta-glucan sy
21	60.0	577	2	D69796	two-component sens
22	60.0	591	2	S65588	ABC-transporter St
23	60.0	610	2	C89939	DnaK protein (limpo
24	60.0	610	2	D84245	DnaK gyrase subunit
25	60.0	639	1	A39135	DNA topoisomerase
26	60.0	671	2	A01666	DNA ligase homolog
27	60.0	671	2	AF1294	DNA ligase homolog
28	60.0	751	2	AB3360	nitrogen regulatio
29	60.0	772	2	A48822	protein-glutamine

30	60.0	1108	2	T17455	translation initia
31	60.0	1114	2	T14351	serine/threonine-s
32	60.0	1374	2	A71724	dna-directed RNA p
33	50.0	56	2	A90156	conserved hypotbet
34	50.0	61	2	B39754	myelin basic prote
35	50.0	62	2	AD1377	hypothetical prote
36	50.0	64	1	Q1BP2L	hypothetical prote
37	50.0	68	2	T02924	acyl carrier prote
38	50.0	82	1	G64370	conserved hypotbet
39	50.0	94	1	IQECAB	integration host f
40	50.0	94	2	A85617	integration host f
41	50.0	94	2	C90753	integration host f
42	50.0	94	2	AG0135	conserved hypotbet
43	50.0	95	2	T49005	hypothetical prote
44	50.0	97	2	S72857	hypothetical prote
45	50.0	99	2	H82572	RNA polymerase ome

ALIGNMENTS

RESULT 1

T05516
hypothetical protein F13M23.90 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 26-Aug-1999
C/Accession: T05516
R/Bevan, M.; Wedler, H.; Wedler, E.; Wandut, R.; Hohenseel, J.; Mewes, H.W.; Mayer, K.F.J.
submitted to the Protein Sequence Database, February 1999
A/Reference number: Z15419
A/Accession: T05516
A/Molecule type: DNA
A/Residues: 1-157

A/Cross-references: EMBL:AL035523
A/Experimental source: cultivar Columbia; BAC clone F13M23
C/Genetics:
A/Map position: 4
A/Note: F13M23.90
C/Superfamily: Arabidopsis hypothetical protein F13M23.90

Query Match
Best Local Similarity 60.0%; Score 6; DB 2; Length 157;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
Db 39 ITVEDA 44

RESULT 2

E82459
probable acetyltransferase VCA0436 [imported] - Vibrio cholerae (strain N16961 serogroup
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C/Accession: E82459
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
charadon, D.; Ermolaeva, M.D.; Vamathevan, J.; Basse, S.; Qin, H.; Dragoti, I.; Sellers, P
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: AB2035; MUID:20406833; PMID:10952301
A/Accession: E82459
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-171 <HRI>
A/Cross-references: GB:AE004377; GB:AE003853; NID:99657831; PIDN:AAF96342.1; GSPDB:GN001;
A/Experimental source: serogroup O1, strain N16961, biotype El Tor
C/Genetics:
A/Gene: VCA0436
A/Map position: 2

Query Match
Best Local Similarity 60.0%; Score 6; DB 2; Length 171;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
|||||
Db 11 TVEDAE 16

RESULT 3
E87299
hypothetical protein CC0406 [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Dec-2002
C/Accession: E87299
R/Niemann, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kjolseth, N.J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: E87299
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-187 <STO>
A/Cross-references: GB:AE005673; NID:913421567; PIDN:AAK2393.1; GSPDB:GN00148
C/Genetics:
A:Gene: CC0406
C:Superfamily: uncharacterized conserved protein

Query Match 60.0%; Score 6; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
|||||
Db 87 ITVEDA 92

RESULT 4

F82451
probable acetyltransferase VCA0505 [imported] - Vibrio cholerae (strain N16961 serogroup C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Esmolaeva, M.D.; Vamshayan, J.; Bess, S.; Qin, H.; Dragoi, I.; Sellers, F.L.R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: F82451
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-197 <HEI>
A/Cross-references: GB:AE004381; GB:AE003853; NID:99657902; PIDN:AAF6408.1; GSPDB:GN001
C/Genetics:
A:Experimental source: serogroup O1, strain N16961, biotype El Tor
A:Gene: VCA0505
A:Map position: 2

Query Match 60.0%; Score 6; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
|||||
Db 37 TVEDAE 42

RESULT 5

JQ1473
pancreatic elastase (EC 3.4.21.36) IV precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 22-Jun-1999

C/Accession: JQ1473; S23783
R/Kang, J.; Wiegand, U.; Mueller-Hill, B. Gene 110, 181-187, 1992
A/Title: Identification of cDNAs encoding two novel rat pancreatic serine proteases.

A/Reference number: JQ1471; MUID:92165057; PMID:1537555
A/Accession: JQ1473
A/Molecule type: mRNA
A/Residues: 1-268 <KAN>
A/Cross-references: EMBL:X59014; NID:956090; PIDN:CAA41753.1; PID:956091
C/Superfamily: trypsin, trypsin homology
C/Keywords: hydrolase; serine proteinase; zymogen
F/1-16/Domain: signal sequence #status predicted <SIG>
F/17-29/Domain: activation peptide #status predicted <ACT>
F/30-268/Product: pancreatic elastase IV #status predicted <MAT>
F/74,121,216/Active site: His, Asp, Ser #status predicted

Query Match 60.0%; Score 6; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
|||||
Db 92 TVEDAE 97

RESULT 6

F83331
Probable transcription regulator PA2511 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: F83331
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: F83331
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-333 <STO>
A/Cross-references: GB:AE004678; GB:AE004091; NID:9948560; PIDN:AAG05899.1; GSPDB:GN001
C/Genetics:
A:Experimental source: strain PA01
A:Gene: PA2511

Query Match 60.0%; Score 6; DB 2; Length 333;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
|||||
Db 71 ITVEDA 76

RESULT 7

G71684
viral protein (viral) RP292 - Rickettsia prowazekii
C/Species: Rickettsia prowazekii
C/Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C/Accession: G71684
R/Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Almark, U. Nature 396, 133-140, 1998
A/Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A/Reference number: A71630; MUID:99039499; PMID:9823893
A/Accession: G71684
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-334 <AND>
A/Cross-references: GB:AJ235271; GB:AJ235269; NID:93868717; PIDN:CAA14753.1; PID:9386085
A:Experimental source: strain Madrid E
C/Genetics:

A:Gene: virB11, RP292
C:Superfamily: tumor-inducing plasmid pTIC58 virB11 protein

Query Match 60.0%; Score 6; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
|||||

Db 191 ITVEDA 196

RESULT 8

F97748
virB11 protein [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001

C:Accession: F97748

R:Gata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rd
Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: F97748

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-334 <STO>

A:Cross-references: GB:AE006914; PIDN:ALU02928.1; PID:G15619456; GSPDB:GN00173
C:Genetics:

A:Gene: virB11

C:Superfamily: tumor-inducing plasmid pTIC58 virB11 protein

Query Match 60.0%; Score 6; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
|||||

Db 191 ITVEDA 196

RESULT 9

T02714
hypothetical protein At2g03120 [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein T18E12.21

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001

C:Accession: T02714; E84444

R:Roundley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rot
submitted to the EMBL Data Library, September 1998

A:Description: Arabidopsis thaliana chromosome II BAC T18E12 genomic sequence.
A:Reference number: Z14702

A:Accession: T02714

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-344 <ROU>

A:Cross-references: EMBL:AC005313; NID:G3548797; PID:G3548818

A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.;
euse, D.; Nieaman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84444

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-344 <STO>

A:Cross-references: GB:AE002093; NID:G3548818; PIDN:AC34490.1; GSPDB:GN00139
C:Genetics:

A:Gene: At2g03120, T18E12.21

A:Map position: 2
A:introns: 14/1; 52/3; 103/2; 132/1; 174/3; 191/3; 220/3; 244/1; 289/3; 315/3

Query Match 60.0%; Score 6; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIV 6
|||||

Db 33 TACTIV 38

RESULT 10

E83920
hypothetical protein BH2165 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: E83920

R:Takant, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: E83920

A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-348 <STO>

A:Cross-references: GB:AP001514; GB:BA000004; NID:G10174613; PIDN:BA05884.1; GSPDB:GN001

A:Experimental source: strain C-125
C:Genetics:

A:Gene: BH2165

Query Match 60.0%; Score 6; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
|||||

Db 104 ITVEDA 109

RESULT 11

T24744
hypothetical protein T09E11.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C:Accession: T24744

R:McLay, K.
submitted to the EMBL Data Library, October 1996

A:Reference number: Z19930

A:Accession: T24744

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-349 <WTL>

A:Cross-references: EMBL:Z81147; PIDN:CAB03538.1; GSPDB:GN00019; CESP:T09E11.11

A:Experimental source: clone T09E11
C:Genetics:

A:Gene: CESP:T09E11.11

A:Map position: 1
A:introns: 66/1; 103/1; 124/1; 191/1; 227/1; 285/3; 315/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T09E5.1

Query Match 60.0%; Score 6; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
|||||

Db 284 ITVEDA 289

RESULT 12
A47211
protein kinase ERK (EC 2.7.1.1-) CEK1 - yeast (Candida albicans) (fragment)

C:Species: Candida albicans
C>Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 10-Jul-1998
C:Accession: A47211

R:Whiteway, M.; Dignard, D.; Thomas, D.Y.
 Proc. Natl. Acad. Sci. U.S.A. 89, 9410-9414, 1992
 A:Title: Dominant negative selection of heterologous genes: isolation of *Candida albicans*
 A:Reference number: A47211; MUID:93028473; PMID:1408649
 A:Accession: A47211
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-362 <WHI>
 A:Note: sequence extracted from NCBI backbone (NCBI:115747)
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
 F:11-316/Domain: protein kinase homology <KIN>
 F:19-27/Region: protein kinase ATP-binding motif

Query Match 60.0%; Score 6; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9
 |||||
 Db 305 ITVEDA 310

RESULT 13
 A60041
 C22/calmodulin-dependent protein kinase (EC 2.7.1.123) ERK1 - Chinese hamster (fragment
 N:Alternate names: extracellular signal-regulated kinase 1 (ERK1); MAP kinase; microtubuli
 C:Species: *Cricetulus griseus* (Chinese hamster)
 C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 17-Mar-1999
 C:Accession: A60041
 R:Meleche, S.; Pages, G.; Pouyssegur, J.
 Mol. Biol. Cell 3, 63-71, 1992
 A:Title: Functional expression and growth factor activation of an epitope-tagged p44 mit
 A:Reference number: A60041; MUID:92199340; PMID:1372523
 A:Accession: A60041
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-369 <MEU>
 C:Comment: This serine/threonine kinase is activated in response to extracellular stimuli
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: ATP; calmodulin binding; phosphoprotein; phosphotransferase; protein kinase
 F:30-320/Domain: protein kinase homology <KIN>
 F:38-46/Region: protein kinase ATP-binding motif

Query Match 60.0%; Score 6; DB 2; Length 369;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9
 |||||
 Db 309 ITVEDA 314

RESULT 14
 S51320
 mitogen-activated protein kinase 6 (EC 2.7.1.1) - common tobacco
 N:Alternate names: serine/threonine-specific protein kinase p43Nt66
 C:Species: *Nicotiana tabacum* (common tobacco)
 C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
 C:Accession: S68189; S51320
 R:Wilson, C.; Anglmayer, R.; Vicente, O.; Heberle-Bors, E.
 Eur. J. Biochem. 233, 249-257, 1995
 A:Title: Molecular cloning, functional expression in *Escherichia coli*, and characterizat
 A:Reference number: S68189; MUID:96061956; PMID:7588752
 A:Accession: S68189
 A:Molecule type: mRNA
 A:Residues: 1-371 <W12>
 A:Cross-references: EMBL:X83879; NID:9634067; PIDN:CAA6760.1; PID:9634068
 C:Genetics:
 A:Gene: nt66
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F:35-324/Domain: protein kinase homology <KIN>

F:44-52/Region: protein kinase ATP-binding motif

Query Match 60.0%; Score 6; DB 2; Length 371;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9
 |||||
 Db 313 ITVEDA 318

RESULT 15
 S48123
 mitogen-activated protein kinase 7 (EC 2.7.1.1) - alfalfa
 N:Alternate names: MAP kinase; MsK7 kinase
 C:Species: *Medicago sativa* (alfalfa)
 C:Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
 C:Accession: S48123; S59949
 R:Jonak, C.; Pay, A.; Boegre, L.; Hirt, H.; Heberle-Bors, E.
 Plant J. 3, 611-617, 1993
 A:Title: The plant homologue of MAP kinase is expressed in a cell cycle-dependent and ori
 A:Reference number: S48123; MUID:94035163; PMID:8220466
 A:Accession: S48123
 A:Molecule type: mRNA
 A:Residues: 1-387 <JON>
 A:Cross-references: EMBL:X66469; NID:9298018; PIDN:CAA47099.1; PID:9298019
 R:Duerr, B.; Gawienowski, M.; Ropp, T.; Jacobs, T.
 Plant Cell 5, 87-96, 1993
 A:Title: MsERK1: a mitogen-activated protein kinase from a flowering plant.
 A:Reference number: S59949; MUID:93177216; PMID:8439746
 A:Accession: S59949
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-387 <DUB>
 A:Cross-references: EMBL:L07042; NID:9289124; PIDN:AAA41548.1; PID:9289125
 C:Genetics:
 A:Gene: MsK7
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: ATP; phosphotransferase; protein kinase
 F:52-340/Domain: protein kinase homology <KIN>
 F:61-69/Region: protein kinase ATP-binding motif

Query Match 60.0%; Score 6; DB 2; Length 387;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9
 |||||
 Db 329 ITVEDA 334

Search completed: December 22, 2003, 11:57:51
 Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:52:00 ; Search time 10 Seconds
(without alignments)
47.027 Million cell updates/sec

Title: US-09-745-763-36_COPY_231_240
Perfect score: 10
Sequence: 1 TACTIVEDAE 10

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues
Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	60.0	310	1 FMT_STRCO	Q910y6 streptomyc
2	60.0	371	1 NTF6_TOBAC	Q40531 nicotiana t
3	60.0	387	1 NMK1_MEDSA	Q07176 medicago sa
4	60.0	393	1 NTF4_TOBAC	Q40532 nicotiana t
5	60.0	417	1 ERK1_CANAL	P28859 candida alb
6	60.0	424	1 MS11_ARATH	Q22467 arabidopsis
7	60.0	433	1 KKO1_YEAST	P36005 saccharomyc
8	60.0	532	1 G61_NEUCR	P38678 neurospora
9	60.0	610	1 DNK1_STAMU	Q98177 staphylococ
10	60.0	610	1 DNK1_STAMU	P45554 staphylococ
11	60.0	639	1 GYRB_HAISSO	P21558 haloferrax s
12	60.0	772	1 ANNU_SCHAM	P52183 schistosom
13	60.0	792	1 CADB_CHICK	Q93131 gallus galli
14	60.0	1108	1 E2K3_RAT	Q92121 rattus norv
15	60.0	1114	1 E2K3_MOUSE	Q92125 mus musculu
16	60.0	1115	1 E2K3_HUMAN	Q92125 homo sapien
17	60.0	1373	1 RPOB_RICMA	Q92143 rickettsia
18	60.0	1374	1 RPOB_RICMA	Q52271 rickettsia
19	50.0	64	1 Y64_LAMB	P03773 bacterioph
20	50.0	67	1 RPO2_PALSO	Q58x90 ralteoncia s
21	50.0	82	1 Y567_METUA	Q57987 methanococ
22	50.0	94	1 IHPB_ECOLI	P08756 escherichia
23	50.0	94	1 Y803_YERPE	Q84112 yersinia pe
24	50.0	99	1 RPO2_XANAC	Q84141 xanthomonas
25	50.0	99	1 RPO2_XYLEA	Q92d77 xyloia fas
26	50.0	102	1 TMOD_PSEME	Q00459 pseudomonas
27	50.0	106	1 YX83_METAC	Q84160 metanococc
28	50.0	108	1 GLPE_SALTY	Q84163 salmoneilla
29	50.0	109	1 VNS2_CVHOC	Q04853 human cocon
30	50.0	121	1 G1N2_METTM	Q50787 methanobac
31	50.0	129	1 ACP2_HORVU	P08817 hordeum vul
32	50.0	132	1 ACP3_HORVU	P15543 hordeum vul
33	50.0	132	1 R58_CORGL	Q84163 corynebacte

34	5	50.0	140	1 SP22_CLOPE	Q8415 clostridium
35	5	50.0	140	1 VMTG_LAMB	P03773 bacterioph
36	5	50.0	146	1 STM2_MOUSE	P37093 mus musculu
37	5	50.0	157	1 YCP6_YEAST	P25130 saccharomyc
38	5	50.0	167	1 SRP_PHAVU	Q41112 phaseolus v
39	5	50.0	175	1 AXIK_ARATH	Q24410 arabidopsis
40	5	50.0	186	1 VNS3_JCDNV	Q90056 junonia coe
41	5	50.0	187	1 EFP_TREPA	Q83537 treponema p
42	5	50.0	194	1 GRPE_BACHD	Q94d73 bacillus ha
43	5	50.0	201	1 ADEN_ADEO7	Q83906 ovine adeno
44	5	50.0	201	1 TMUO_BSPBP	P06650 bacterioph
45	5	50.0	214	1 YA04_METUA	Q58410 methanococc

ALIGNMENTS

RESULT 1
FMT_STRCO STANDARD; PRT; 310 AA.
ID FMT_STRCO
AC Q910y6:
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Methionyl-tRNA formyltransferase (EC 2.1.2.9).
DS FMT OR SC01473 OR SCL6.30C.
GN Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STPAIN-A3(2) / M145;
RX MEDLINE=21996410; Pubmed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.B., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornaby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajadream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)." (2002).
RL Nature 417:141-147(2002).
RC -1- FUNCTION: Modify the free amino group of the aminoacyl moiety of
methionyl-tRNA(fMet). The formyl group appears to play a dual role
in the initiator identity of N-formylmethionyl-tRNA by: (i)
promoting its recognition by IF2 and (ii) impairing its binding to
EF-Tu-GTP (By similarity).
CC -1- CATALYTIC ACTIVITY: 10-formyltetrahydrofolate + L-methionyl-
tRNA(fMet) + H(2)O = tetrahydrofolate + N-formylmethionyl-
tRNA(fMet).
CC -1- SIMILARITY: Belongs to the fmt family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
or send an email to license@ebi.ac.uk).

CC EMBL; AL39109; CAB76895.1; -
DR HSSP; P23882; 1FMT.
DR HAMAP; MF_00182; -; 1.
DR InterPro; IPR005794; Fmt.
DR InterPro; IPR005793; Formyl_trans_C.
DR InterPro; IPR002376; Formyl_transf.
DR Pfam; PF02911; formyl_trans_C; 1.
DR Pfam; PF00551; formyl_transf; 1.

DR TIGRFAMs: TIGR00460; fnc; 1.
 DR PROSITE; PS00373; GART; FALSE NEG.
 KW Transferrase; Methyltransferase; Protein biosynthesis;
 KW Complete Proteome.
 FT BINDING 110 113 TETRAHYDROFOLATE (THF) (BY SIMILARITY).
 SQ SEQUENCE 310 AA; 32639 MW; A2BA4098D6F21C2E CRC64;
 Query Match 60.0%; Score 6; DB 1; Length 310;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 4 ITVEDA 9
 Db 205 ITVEDA 210

RESULT 2
 NTF6_TOBAC STANDARD; PRT; 371 AA.
 ID NTF6_TOBAC
 AC Q40531;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Mitogen-activated protein kinase homolog NTF6 (EC 2.7.1.-) (P43).
 GN NTF6.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asterids; Lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Petit Havana SRI;
 RX MEDLINE=96061956; PubMed=7588752;
 RA Wilson C., Angimayer R., Vicente O., Heberle-Bors E.,
 RT "Molecular cloning, functional expression in *Escherichia coli*, and
 RT characterization of multiple mitogen-activated-protein kinases from
 RT tobacco.";
 RL Eur. J. Biochem. 233:249-257(1995).

CC -1- ENZYME REGULATION: Activated by tyrosine and threonine
 CC phosphorylation (By similarity).
 CC -1- PTM: VERY LOW AUTOPHOSPHORYLATION, ALTHOUGH DRAMATICALLY INCREASED
 CC WHEN MN(2+) IS ADDED TO THE REACTION INSTEAD OF MG(2+).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X63879; CAA58760.1; -.
 DR PIR; S68189; S51320.
 DR HSSP; Q16539; 1MFC.
 DR InterPro; IPR003527; MAP_kin.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS01351; MAPK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Multigene family;
 FT DOMAIN 38 324 PROTEIN KINASE.
 FT NP_BIND 44 52 ATP (BY SIMILARITY).
 FT BINDING 67 67 ATP (BY SIMILARITY).
 FT ACT_SITE 164 164 BY SIMILARITY.

FT MOD_RES 196 196 PHOSPHORYLATION (ACTIVATES THE KINASE)
 FT FT (BY SIMILARITY).
 FT MOD_RES 198 198 PHOSPHORYLATION (ACTIVATES THE KINASE)
 FT FT (BY SIMILARITY).
 SQ SEQUENCE 371 AA; 42741 MW; 4D97C41AC203C272 CRC64;
 Query Match 60.0%; Score 6; DB 1; Length 371;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 4 ITVEDA 9
 Db 313 ITVEDA 318

RESULT 3
 MKK1_MEDSA STANDARD; PRT; 387 AA.
 ID MKK1_MEDSA
 AC Q07176;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Mitogen-activated protein kinase homolog MKK1 (EC 2.7.1.-) (MAP kinase
 DE MSK7) (MAP kinase ERK1).
 GN MKK1 OR MSK7 OR ERK1.
 OS Medicago sativa (Alfalfa).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eubosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.
 OX NCBI_TaxID=3879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=94035163; PubMed=8220466;
 RA Jonak C., Pay A., Boegre U., Hirt H., Heberle-Bors E.,
 RT "The plant homologue of MAP kinase is expressed in a cell cycle-
 RT dependent and organ-specific manner.";
 RL Plant J. 3:611-617(1993).

CC -1- FUNCTION: MAY PLAY A ROLE IN THE MITOGENIC INDUCTION OF SYMBIOTIC
 CC ROOT NODULES ON ALFALFA BY RHIZOBIUM SIGMAU MOLECULUS.
 CC -1- ENZYME REGULATION: Activated by tyrosine and threonine
 CC phosphorylation (By similarity).
 CC -1- TISSUE SPECIFICITY: ROOTS AND STEMS.
 CC -1- PTM: Autophosphorylated.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X66469; CAA47099.1; -.
 DR EMBL; U07042; AAB41548.1; -.
 DR PIR; S48123; S48123.
 DR HSSP; P27703; 1ERK.
 DR InterPro; IPR003527; MAP_kin.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS01351; MAPK; 1.

[illegible]

OY	4	ITVEDA	9
DB	360	ITVEDA	365

RESULT 6

MS11_ARATH			
ID	MS11_ARATH	STANDARD:	PRT; 424 AA.
AC	O22457;		
DT	15-JUL-1998	(Rel. 36, Created)	
DT	15-JUL-1998	(Rel. 36, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	WD-40 repeat protein MS11.		
GN	MS11 OR AT5G58220 OR MCK7.10.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_TaxId=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=97480095; PubMed=9338962;		
RA	Ach R.A., Taranto P., Gruijssem W.;		
RT	"A conserved family of WD-40 proteins binds to the retinoblastoma		
RT	protein in both plants and animals."		
RL	Plant Cell 9:1595-1606(1997).		
RN	[2]		

```

RP SEQUENCE FROM N.A.
RC STAIN=CV, Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
  Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
  features of the regions of 3,076,755 bp covered by sixty P1 and TAC
  clones.";
RL DNA Res. 7:31-63(2000).
CC -1- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -1- SIMILARITY: Contains 6 WD repeats.
CC -1- SIMILARITY: BELONGS TO THE RBAP46/RBAP48/MS11 FAMILY OF WD-REPEAT
  PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC DR EMBL, AF016846; AAB70242.1; -
CC DR EMBL, AB019228; AAB6914.1; -
CC DR InterPro: IPR001680; WD40.
CC DR Pfam: PF00400; WD40; 5
CC DR SMART, SM00320; WD40; 6.
CC DR PROSITE, PS00678; WD_REPEATS_1; 1.
CC DR PROSITE, PS50082; WD_REPEATS_2; 5.
CC DR PROSITE, PS50294; WD_REPEATS_REGION; 1.
CC KW Repeat; WD repeat; Nuclear protein.
CC FT REPEAT 123 163 WD 1.
CC FT REPEAT 176 216 WD 2.
CC FT REPEAT 225 265 WD 3.
CC FT REPEAT 271 311 WD 4.
CC FT REPEAT 315 355 WD 5.
CC FT REPEAT 372 412 WD 6.
CC SQ SEQUENCE 424 AA; 48194 MW; 367D89C861ECADF CRC64;

Query Match 60.0%; Score 6; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
    |||||
Db 355 TVEDAE 360

RESULT 7
REQL_YEAST
ID_KKQ1_YEAST STANDARD; PRT; 433 AA.
AC P36005;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable serine/threonine-protein kinase YKL161C (EC 2.7.1.-).
GN YKL161C OR YKL615.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=94378720; PubMed=8091859;
RA Vandenbol M., Bolle P.-A., Dion C., Portetelle D., Hilger F.;
RT "DNA sequencing of a 36.2 kb fragment located between the FAS1 and
  LAP loci of chromosome XI of Saccharomyces cerevisiae.";
RL Yeast 10:535-540(1994).
CC -1- SIMILARITY: BELONGS TO THE SBR/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z26877; CA81493.1; -
DR EMBL; Z28161; CA82003.1; -
DR PIR; S37790; S37790.
DR HSSP; P27703; IGOL.
DR SGD; S0001644; YKL161C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 23 318 PROTEIN_KINASE.
FT NP_BIND 29 37 ATP (BY SIMILARITY).
FT BINDING 55 55 ATP (BY SIMILARITY).
FT ACT_SITE 153 153 BY SIMILARITY.
SQ SEQUENCE 433 AA; 49632 MW; BC27B76B7D95822 CRC64;

Query Match 60.0%; Score 6; DB 1; Length 433;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
Db 307 ITVEDA 312

RESULT 8
GSI_NEUCR STANDARD; PRT; 532 AA.
AC P36678;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucan synthase-1 (EC 2.4.1.34) (1,3-beta-glucan synthase)
DE (UDP-glucose-1,3-beta-D-glucan glucosyltransferase).
GN GS-1 OR B13020.130.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=95023939; PubMed=7937796;
RA Enderlin C.S., Selitrenikoff C.P.;
RT "Cloning and characterization of a Neurospora crassa gene required
RT for (1,3) beta-glucan synthase activity and cell wall formation.";
RT Proc. Natl. Acad. Sci. U.S.A. 91:9500-9504(1994).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX PubMed=12655011;
RA Mannhaupt G., Montreone C., Haase D., Mewes H.-W., Aign V.,
RA Hodelsel J.D., Fartmann B., Nyakatura G., Kempken F., Mäler J.,
RA Schulte U.;
RT "What's in the genome of a filamentous fungus? Analysis of the
RT Neurospora genome sequence.";
RT Nucleic Acids Res. 31:1944-1954(2003).
RL -1- FUNCTION: INVOLVED IN (1,3)BETA-GLUCAN SYNTHASE ACTIVITY AND
CC CELL-WALL FORMATION.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + ((1,3)-beta-D-glucosyl) (N) = UDP
CC + ((1,3)-beta-D-glucosyl) (N+1).
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U09275; AAA50440.1; -
DR EMBL; AL451015; CAC18203.1; -
DR PIR; T46649; T46649.
DR Transferase; Glycosyltransferase.
SQ SEQUENCE 532 AA; 58939 MW; A36C00CA96F2F21B CRC64;

Query Match 60.0%; Score 6; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
Db 520 TVEDAE 525

RESULT 9
DNAX_STAM STANDARD; PRT; 610 AA.
AC O9977;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein dnax (Heat shock protein 70) (Heat shock 70 kDa
DE protein) (HSP70).
GN DNAX OR SAV1580 OR SA1409 OR MW1532.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MU50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iino T., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hasegawa A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hiraoka H., Kohara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashta A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Amano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
RN [2]
RP -1- FUNCTION: Acts as a chaperone (By similarity).
CC -1- INDUCTION: By stress conditions e.g. heat shock (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).


```

CC -----
DR EMBL: AP003362; BAB57742.1; -
DR EMBL: AP003134; BAB42672.1; -
DR EMBL: AP004827; BAB95397.1; -
DR PIR: C89939; C89939.
DR HSSP: P04475; 1D64.
DR HAMAP: MF_00332; -; 1.
DR InterPro: IPR001023; Hsp70.
DR Pfam: PF000012; HSP70_1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PRODOM: PD000089; HSP70_1.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation;
KW Complete proteome.
KW MOD_RES 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 610 AA; 66361 MW; B71645C36D26AED CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 610;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
DB 595 TVEDAE 600

RESULT 10
DNAX_STAUA STANDARD; PRT; 610 AA.
ID DNAX_STAUA STANDARD; PRT; 610 AA.
AC P45554;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein dnax (Heat shock protein 70) (Heat shock 70 kDa
DE protein) (HSP70).
GN DNAX.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94321356; PubMed=8045913;
RT "Molecular cloning of two new heat shock genes related to the hsp70
RT genes in Staphylococcus aureus."
RL J. Bacteriol. 176:4779-4783(1994).
CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC -1- INDUCTION: BY stress conditions e.g. heat shock (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D30690; BAA06359.1; -
DR HSSP: P04475; 1D64.
DR HAMAP: MF_00332; -; 1.
DR InterPro: IPR001023; Hsp70.
DR Pfam: PF000012; HSP70_1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PRODOM: PD000089; HSP70_1.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
KW MOD_RES 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

```

```

SQ SEQUENCE 610 AA; 66347 MW; B34745C36D26AED CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 610;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
DB 595 TVEDAE 600

RESULT 11
GRRB_HALSQ STANDARD; PRT; 639 AA.
ID GRRB_HALSQ STANDARD; PRT; 639 AA.
AC P21558;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA gyrase subunit B (EC 5.99.1.3).
GN GYRB.
OS Haloflex sp. (strain Aa 2.2).
OC Archaea; Euryarchaeota; Halobacteriales;
OC Halobacteriaceae; Haloflex.
OX NCBI_TaxID=2254;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91100352; PubMed=1846146;
RA Holmes M.L., Dyall-Smith M.L.;
RT "Mutations in DNA gyrase result in novobiocin resistance in
RT halophilic archaeobacteria."
RL J. Bacteriol. 173:642-648(1991).
CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS.
CC -1- MISCELLANEOUS: MUTATIONS IN DNA GYRASE RESULT IN NOVOBIOCIN
CC RESISTANCE IN HALOPHILIC ARCHAEABACTERIA.
CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M38373; AAB09605.1; -
DR HSSP: P06982; 1A06.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR002288; DNA_GYRB_C.
DR InterPro: IPR000565; DNA_GYRB.
DR InterPro: IPR001241; DNA_topoisII.
DR InterPro: IPR006171; Toprim_dom.
DR Pfam: PF00204; DNA_gyraseB_1.
DR Pfam: PF00986; DNA_gyraseB_C_1.
DR Pfam: PF02518; HATPase_C_1.
DR Pfam: PF01751; Toprim_1.
DR PRINTS: PR00418; TP12FAMILY.
DR PRODOM: PD148633; DNA_gyraseB_C_1.
DR SMART: SM00387; HATPase_C_1.
DR SMART: SM00433; TOP2c_1.
DR TIGRFAMs: TIGR01059; gyrb_1.
DR PROSITE: PS00177; TOPOISOMERASE II; 1.
KW Topoisomerase; Isomerase; ATP-binding.
SQ SEQUENCE 639 AA; 71138 MW; 691315972C5E7506 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 639;

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TWEDA 9
|||||
Db 600 ITVEDA 605

RESULT 12

ANNU SCHAM STANDARD; PRT: 772 AA.

AC PS2183;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Annullin (Protein-glutamyl gamma-glutamyltransferase) (EC 2.3.2.13) (Transglutaminase).
OS Schistocerca americana (American grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota; Neoptera; Orthopteroidea; Orthoptera; Caellifera; Acridomorpha; Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
NCBI_Taxid=7009;
[1]
RN SEQUENCE FROM N.A.
RX MBLDLINE=93050772; PubMed=1358727;
RA Singer M.A., Hortsch M., Goodman C.S., Bentley D.;
RT "Annullin, a protein expressed at limb segment boundaries in the grasshopper embryo, is homologous to protein cross-linking transglutaminases".
RT Dev. Biol. 154:143-159(1992).
CC -1- FUNCTION: PARTICIPATES IN MORPHOGENETIC ACTIVITIES OF THE CELLS, MAYBE BY STABILIZING THE MEMBRANE OR SUBCORTICAL STRUCTURES OF CELLS THAT ARE UNDER MECHANICAL STRESS. PROBABLY CATALYZES THE CROSS-LINKING OF PROTEINS AND THE CONJUGATION OF POLYAMINES TO PROTEINS.
CC -1- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N(5)-alkylglutamine + NH(3).
CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: INTRACELLULAR AND PERIPHERALLY ASSOCIATED WITH THE INNER LEAFLET OF THE CELL MEMBRANE, USING A FATTY ACID LINKAGE.
CC -1- TISSUE SPECIFICITY: HAS AN ANNULAR, OR RING-LIKE EXPRESSION PATTERN IN EPITHELIAL ANNULI OF DEVELOPING LIMB SEGMENT BOUNDARY CELLS. IN EMBRYOS, IT IS SEEN IN GASTRULATING CELLS, IN CELLS SURROUNDING RAPIDLY DIVIDING NEUROBLASTS, AND IN MUSCLE PIONEER CELLS INVAGINATING TO FORM APODEMS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION OF THIS PROTEIN IN EMBRYOS AND LIMBS IS ASSOCIATED WITH AREAS UNDERGOING MOVEMENTS, MORPHOGENETIC REARRANGEMENTS, OR RAPID CELL DIVISION. EXPRESSION OF ANNULIN PRECEDES THE FIRST MORPHOLOGICAL SIGNS OF SEGMENTATION IN THE DEVELOPING LIMBS.
CC -1- SIMILARITY: BELONGS TO THE TRANSGLUTAMINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M92291; AAA29806.1; -
CC PIR; A48822; A48822.
CC HSP; P00488; 1GCU.
DR InterPro; IPR001102; Gluttransfg.
DR InterPro; IPR002931; Tnsglutase_like.
DR Pfam; PF01841; Tnsglut_core; 1.
DR Pfam; PF00927; Transglutamin_C; 2.
DR Pfam; PF00868; Transglutamin_N; 1.
DR SMART; SM00460; TGC; 1.
DR PROSITE; PS00547; TRANSGLUTAMINASES; 1.
DR Transferrase; Acyltransferase; Calcium-binding; Lipoprotein; Palmitate.
KW TRANSFERRASE; ACYLTRANSFERASE; CALCIUM-BINDING; LIPOPROTEIN; PALMITATE.
FT LIPOD 4 5.
FT PALMITATE (POTENTIAL).
FT LIPOD 5 5.
FT PALMITATE (POTENTIAL).

FT ACT_SITE 400 400 BY SIMILARITY.
FT ACT_SITE 427 427 BY SIMILARITY.
FT METAL 467 467 CALCIUM (BY SIMILARITY).
FT METAL 469 469 CALCIUM (BY SIMILARITY).
FT METAL 517 517 CALCIUM (BY SIMILARITY).
FT METAL 522 522 CALCIUM (BY SIMILARITY).
SQ SEQUENCE 772 AA; 85941 MW; FA5A3CE6A7CAE394 CRC64;

Query Match 60.0%; Score 6; DB 1; Length 772;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TWEDA 10
|||||
Db 127 ITVEDA 132

RESULT 13

CADB CHICK STANDARD; PRT: 792 AA.

AC 093319;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Cadherin-11 precursor.
GN CDH11.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
NCBI_Taxid=9031;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=White Leghorn;
RA Wei J., Dong X.R., Topouzis S., Zimmer W.E., Broders F., Thiery J.P., Kotteliansky V., Majesky M.W.;
RT "Molecular cloning of chick cadherin 11 and its expression during smooth muscle differentiation and formation of the tunica media.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS. THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE SORTING OF HETEROGENEOUS CELL TYPES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (potential).
CC -1- SIMILARITY: Contains 5 cadherin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF055342; AAC33675.1; -
CC InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR Pfam; PF01049; cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00233; CADHERIN_1; 3.
DR PROSITE; PS50268; CADHERIN_2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 22
FT PROPEP 23 53
FT CHAIN 54 792
FT CHAIN 54 613
FT TRANSMEM 614 634
FT DOMAIN 635 792
FT DOMAIN 54 159
FT DOMAIN 160 268
FT DOMAIN 160 268 CADHERIN 2.

FT DOMAIN 269 383 CADHERIN 3.
FT DOMAIN 384 486 CADHERIN 4.
FT DOMAIN 487 608 CADHERIN 5.
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 792 AA; 87572 MW; 3E3489C6686731AB CRC64;
Query Match 60.0%; Score 6; DB 1; Length 792;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 TVEDA 9
Db 372 TVEDA 377
RESULT 14
E2K3 RAT STANDARD; PRT; 1108 AA.
AC 092121;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eukaryotic translation initiation factor 2-alpha kinase 3 precursor
DE (EC 2.7.1.-) (PKR-like endoplasmic reticulum kinase) (Pancreatic
DE eIF2-alpha kinase).
CN EIF2AK3 OR PERK OR PERK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Pancreatic Islets;
RX MEDLINE=99038253; PubMed=9819435;
RA Shi Y., Vattam K.M., Sood R., An J., Liang J., Stramm L.E., Wek R.C.;
RT "Identification and characterization of pancreatic eukaryotic
RT initiation factor 2 alpha-subunit kinase, PERK, involved in
RT translational control.";
RT Mol. Cell. Biol. 18:7499-7509 (1998).
RN [2]
RN MUTAGENESIS OF LYS-614.
RP MEDLINE=99150360; PubMed=10026192;
RA Shi Y., An J., Liang J., Hayes S.E., Sandusky G.E., Stramm L.E.,
RA Yang N.N.;
RT "Characterization of a mutant pancreatic eIF-2alpha kinase, PERK, and
RT co-localization with somatostatin in islet delta cells.";
RT J. Biol. Chem. 274:5723-5730 (1999).
RN [3]
RP SUBUNIT.
RX MEDLINE=20313073; PubMed=10854322;
RA Bertolotti A., Zhang Y., Hendershot L.M., Harding H.P., Ron D.;
RT "Dynamic interaction of BiP and ER stress transducers in the
RT unfolded-protein response.";
RT Nat. Cell Biol. 2:332-332 (2000).
CC -1- FUNCTION: PHOSPHORYLATES THE ALPHA SUBUNIT OF EUKARYOTIC
CC INACTIVATION-INITIATION FACTOR 2 (EIF2), LEADING TO ITS
CC INACTIVATION AND THUS TO A RAPID REDUCTION OF TRANSLATIONAL
CC INITIATION AND REPRESSION OF GLOBAL PROTEIN SYNTHESIS. SERVES AS A
CC CRITICAL EFFECTOR OF UNFOLDED PROTEIN RESPONSE (UPR)-INDUCED G1
CC GROWTH ARREST DUE TO THE LOSS OF CYCLIN D1 (BY SIMILARITY).
CC -1- ENZYME REGULATION: PERTURBATION IN PROTEIN FOLDING IN THE
CC ENDOPLASMIC RETICULUM (ER) PROMOTES REVERSIBLE DISSOCIATION FROM
CC HSPAS/BIP AND OLIGOMERIZATION, RESULTING IN
CC TRANSAUTOPHOSPHORYLATION AND KINASE ACTIVITY INDUCTION.
CC -1- SUBUNIT: FORMS DIMERS WITH HSPAS/BIP IN RESTING CELLS.
CC OLIGOMERIZES IN ER-STRESSED CELLS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
CC reticulum.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- INDUCTION: By ER stress.
CC -1- DOMAIN: THE LUMENAL DOMAIN SENSES PERTURBATIONS IN PROTEIN FOLDING

CC IN THE ER, PROBABLY THROUGH REVERSIBLE INTERACTION WITH HSPAS/BIP.
CC -1- PTM: AUTOPHOSPHORYLATED.
CC -1- PTM: N-GLYCOSYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC GCM2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF096835; AAC83801.1; -.
DR PIR; T17455; T17455.
DR InterPro; IPR000715; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Serine/threonine-protein kinase; transferase; ATP-binding;
KW Translation regulation; Unfolded protein response;
KW Endoplasmic reticulum; Phosphorylation; Glycoprotein; Signal;
KW Transmembrane.
FT SIGNAL 1 27
FT CHAIN 28 1108
FT DOMAIN 28 506
FT TRANSMEM 507 527
FT DOMAIN 528 1108
FT DOMAIN 585 1069
FT NP_BIND 591 599
FT BINDING 614 614
FT ACT_SITE 929 929
FT DOMAIN 47 50
FT DOMAIN 223 228
FT CARBOHYD 253 253
FT MUTAGEN 614 614
SQ SEQUENCE 1108 AA; 124769 MW; B3716B1FD26ED32B CRC64;
Query Match 60.0%; Score 6; DB 1; Length 1108;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 TVEDAB 10
Db 66 TVEDAB 71
RESULT 15
E2K3 MOUSE STANDARD; PRT; 1114 AA.
AC 0922B5; 09CTK8; 09CWT5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eukaryotic translation initiation factor 2-alpha kinase 3 precursor
DE (EC 2.7.1.-) (PKR-like endoplasmic reticulum kinase) (Pancreatic
DE eIF2-alpha kinase).
CN EIF2AK3 OR PERK OR PERK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-618.
RC STRAIN=NIH Swiss; TISSUE=Fibroblast;
RX MEDLINE=99127894; PubMed=9930704;
RA Harding H.P., Zhang Y., Ron D.;
RT "Translation and protein folding are coupled by an endoplasmic-

RT. reticulum-resident kinase.";
 RL Nature 397:271-274(1999).
 RN [2]
 RP ERRATUM.
 RA Harding H.P., Zhang Y., Ron D.;
 RL Nature 398:90-90(1999).
 RN [3]
 RP SEQUENCE OF 1-152 AND 769-1114 FROM N.A..
 RC STRAIN=C57BL/6J;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ienhi Y.,
 RA Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kaubuka T., Saito R.,
 RA Kodora K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Scriml L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamata M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [4]
 RP SUBUNIT.
 RX MEDLINE=20313073; PubMed=10854322;
 RA Bertolotti A., Zhang Y., Hendershot L.M., Harding H.P., Ron D.;
 RT "Dynamic interaction of BiP and ER stress transducers in the
 RT unfolded-protein response.";
 RL Nat. Cell Biol. 2:326-332(2000).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=20524051; PubMed=11035797;
 RA Brewer J.W., Diehl J.A.;
 RT "PERK mediates cell-cycle exit during the mammalian unfolded protein
 RT response.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12625-12630(2000).
 CC -1- FUNCTION: PHOSPHORYLATES THE ALPHA SUBUNIT OF EUKARYOTIC
 CC INACTIVATION-INITIATION FACTOR 2 (EIF2), LEADING TO ITS
 CC INACTIVATION AND THUS TO A RAPID REDUCTION OF TRANSLATIONAL
 CC INITIATION AND REPRESSION OF GLOBAL PROTEIN SYNTHESIS. SERVES AS A
 CC CRITICAL EFFECTOR OF UNFOLDED PROTEIN RESPONSE (UPR)-INDUCED G1
 CC GROWTH ARREST DUE TO THE LOSS OF CYCLIN D1.
 CC -1- ENZYME REGULATION: PERTURBATION IN PROTEIN FOLDING IN THE
 CC ENDOPLASMIC RETICULUM (ER) PROMOTES REVERSIBLE DISSOCIATION FROM
 CC HSP45/BIP AND OLIGOMERIZATION, RESULTING IN
 CC TRANSAUTOPHOSPHORYLATION AND KINASE ACTIVITY INDUCTION.
 CC -1- SUBUNIT: FORMS DIMERS WITH HSP45/BIP IN RESTING CELLS.
 CC OLIGOMERIZES IN ER-STRESSED CELLS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
 CC reticulum.
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- INDUCTION: By ER stress.
 CC -1- DOMAIN: THE LUMENAL DOMAIN SENSES PERTURBATIONS IN PROTEIN FOLDING
 CC IN THE ER, PROBABLY THROUGH REVERSIBLE INTERACTION WITH HSP45/BIP.
 CC -1- PTM: AUTOPHOSPHORYLATED.
 CC -1- PTM: N-GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC GCN2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 CC DR EMBL: AF076681; AAD03337.1; -;
 CC DR EMBL: AK010387; BAB26908.1; -;
 CC DR EMBL: AK003226; BAB2655.1; -;
 CC DR PIR: T14351; T14351.
 CC DR MCD: MGI:1341830; E1f2ak3.
 CC DR InterPro: IPR000719; Prot kinase.
 CC DR InterPro: IPR002290; Ser Thr kinase.
 CC Pfam: PF00063; Pkinase_1.
 CC DR ProDom: PD000001; Prot kinase; 2.
 CC DR PROSITE: PS00107; PROTEIN KINASE ATP; FALSE NEG.
 CC DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
 CC DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 CC KW Serine/threonine-protein kinase; Transferase; ATP-binding;
 KW Translation regulation; Unfolded protein response;
 KW Endoplasmic reticulum; Phosphorylation; Glycoprotein; Signal;
 KW Transmembrane.
 KW FT SIGNAL 1 28 POTENTIAL.
 KW FT CHAIN 29 1114 EUKARYOTIC TRANSLATION INITIATION FACTOR
 FT 2-ALPHA KINASE 3.
 FT DOMAIN 29 510 POTENTIAL.
 FT TRANSMEM 511 531 LUMENAL (POTENTIAL).
 FT DOMAIN 532 1114 POTENTIAL.
 FT DOMAIN 589 1075 CYTOPLASMIC (POTENTIAL).
 FT NP BIND 595 603 PROTEIN KINASE.
 FT BINDING 618 603 ATP (BY SIMILARITY).
 FT ACT_SITE 935 935 ATP.
 FT DOMAIN 48 51 BY SIMILARITY.
 FT CARBOHYD 254 229 POLY-ALA.
 FT MUTAGEN 618 618 N-LINKED (GLNAC. . .).
 FT CONFLICT 143 152 K-A: LOSS OF ACTIVITY AND
 FT CONFLICT 826 826 AUTOPHOSPHORYLATION.
 FT SEQUENCE 1114 AA; 124681 MW; 65A47D6CDD28046 CRC64;
 SO
 Query Match 60.0%; Score 6; DB 1; Length 1114;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 5 TVEDAE 10
 Db 67 TVEDAE 72
 Search completed: December 22, 2003, 11:56:29
 Job time : 12 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 22, 2003, 15:10:16 ; Search time 358 Seconds
(without alignments)
3559.036 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417
Sequence: 1 MKFLIFAFG3VHLSICSG.....AWMAVSVVADMEEMTPRS 472

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q=/cgn2_1/USPIO_epool_p/US09745763/runat_22122003_113316_25243/app_query.fasta_1.647
-DB=Geneseq_19Jun03 -QMT=faeclap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09745763 @CGN 1.1 396 @runat 22122003 113316 25243 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_19Jun03:*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Length	DB	ID	Description
1	2417	100.0	1851	19	AAV82779
2	2417	100.0	1851	24	ABO92016
3	2417	100.0	1863	21	AAZ98034
4	2417	100.0	1863	22	AAU11647
5	2417	100.0	1863	24	ABK69743
6	2417	100.0	1863	25	ACC50817
7	2417	100.0	1863	25	ABZ71453
8	2417	100.0	1884	21	AAZ58313
9	2417	100.0	1923	21	AAZ98139
10	2417	100.0	2077	25	ACC50603
11	2417	100.0	2077	25	ABZ71331
12	2410	99.7	1767	21	AAA40493
13	2410	99.7	1778	18	AAV02296
14	2402	99.4	1860	22	AAK94491
15	2339.5	96.8	1895	22	AAH99703
16	2259.5	93.5	1784	23	AAZ73592
17	2116	87.5	1778	24	ABK62598
18	1414	58.5	895	21	AAA44369
19	1338	55.4	1134	21	AAZ98101
20	1338	55.4	1134	22	AAU11714
21	1338	55.4	1134	24	ABK69810
22	1338	55.4	1134	25	ACC50818
23	1338	55.4	1134	25	ABZ71454
24	870	36.0	609	20	AAV87487
25	771	31.9	642	22	AAK92181
26	491	20.3	462	22	AAK93447
27	491	20.3	462	18	AAZ79398
28	491	20.3	462	18	AAZ80600
29	491	20.3	462	18	AAV02139
30	479	19.8	314	20	AAV63369
31	397	16.4	441	22	AAK56831
32	351	14.5	317	20	AAK40587
33	229.5	9.5	1491	20	AAK00070
34	229.5	9.5	1491	20	AAV82521
35	229.5	9.5	1491	20	AAK00018
36	226.5	9.4	1593	22	AAK13324
37	225.5	9.3	1640	21	AAZ45188
38	190	7.9	1346	24	ABK75046
39	188	7.8	3110	21	AAZ58312
40	188	7.8	3171	22	AAZ93781
41	188	7.8	3771	22	AAH98667
42	183	7.6	492	22	AAK92932
43	172.5	7.1	2271	21	AAK94453
44	169	7.0	1179	17	AAZ13568
45	169	7.0	2256	21	AAA09458

ALIGNMENTS

RESULT 1
AAV82779
AAV82779 standard; cDNA, 1851 BP.
AAV82779;
25-FEB-1999 (first entry)
Clone bu45_2 isolated from human adult placenta cDNA library.
Secreted protein; nutritional activity; immune stimulating; vaccine;
suppressing activity; haematopoiesis regulating activity;
tissue growth activity; activin; inhibin activity; chemotaxis;
chemokine activity; haemostasis; thrombolytic activity; receptor;
ligand; anti-inflammatory; cadherin; tumour invasion suppressor;
tumour inhibition; gene therapy; ds.

OS Homo sapiens.
 XX MN MO842739-A2.
 XX PD 01-OCT-1998.
 XX PF 20-MAR-1998; 98MO-US05653.
 XX PR 19-MAR-1998; 98US-0044466.
 XX PR 21-MAR-1997; 97US-0822167.
 XX PA (GENY) GENETICS INST INC.
 XX PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie LA, Spaulding V, Treacy M;
 XX DR WPI; 1998-609890/51.
 XX DR P-PSDB; AAW85456.
 PT New polynucleotides encoding secreted human proteins - derived from
 PT human foetal brain, adult brain, foetal kidney, placenta or adult
 PT pineal gland cDNA libraries.
 XX PS Claim 14; Page 69-70; 113pp; English.
 XX CC The present sequence encodes a secreted protein. The polynucleotide and
 CC secreted protein are predicted to have biological activities which would
 CC make them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is given.
 CC Suggested activities include nutritional activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity (no data is
 CC given in the specification to support these activities). The
 CC polynucleotide is also stated to be useful for gene therapy.
 XX SQ Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,096-236 Length: 1851
 Score: 2417.00 Matches: 472
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 19 Gaps: 0
 US-09-745-763-36 (1-472) x AAW82779 (1-1851)
 QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuGlySerGly 20
 DB 99 ATGAATATCTTATCTTCGATTTTCGGTGTGTCACCTTTATCCCTGTGCTGGG 158
 QY 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40
 DB 159 AAAGCTATATGCAAGATGCAATCTCTTAAGAGCACTTTGAAGAAATTAAGAAATTA 218
 QY 41 AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValIleGlyLysAlaGln 60
 DB 219 GCCACGTGGAGATGTTCTTAAGCAATCATCACTAGCTGTTATGTAAGCCCGG 278
 QY 61 AsnArgSerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80
 DB 279 AACAGATCTATAGAGGATGGACACTTCGTGTGATACCTGTGACCCAGACTGAGTGC 338
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
 DB 339 TCCAGAACTTGAAGAAATCCATCCAAATTAATGACCAAACTGCAGCAAGATGGGCTG 398
 QY 101 GluLysValHisLeuGluProValArgIleProHisIleProGluArgGlyGluGluSerAla 120
 DB 399 GAGAAAGTTCACTGGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAAGATCAGCT 458

QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGluGlySerSerIleGly 140
 DB 459 GTGATCTGAGAGCCAAAGATTCATTAAGATAGCCATCTGGGCTTGGCAAGACATGGG 518
 QY 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
 DB 519 ACTCTCCAGAAAGGACATTACAGCAGAAATTCGTGGTGTGACCTCTTCGATGAATCAG 578
 QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
 DB 579 AGAAGGCTCCAGAGAGAGGAGAAATGTTGTTATTAACCAACTTACATCACTAC 638
 QY 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaLysValAlaLeu 200
 DB 639 TCAAGAGCGGTGCAATTCAGAACCGGCGGTGAGCTGCACAGTGGGGCTTGG 698
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleGln 220
 DB 699 GCATCTCTCATTCGATCCGTGGCTCTCTCCATCTACAGTCTCAGACAGTATTCAG 758
 QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 DB 759 GAATACCAAGATGGCGGCCCAAAATTCACACGCTGATTTACGGTGGAGATGACAGAA 818
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 DB 819 ATGATGTCAGAAATGCGCTTCATGAGATCAAAATGTCATTCAGCTTAAGATGGGCA 878
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
 DB 879 AAAGCTACCAAGATGATGATCTCTCAACACTGTACAGAGATCAGTGGAGCAAAATAT 938
 QY 281 ProGluGluValIleValLeuValSerGlyHisIleAspSerTyrAspValGluGlnGlyAla 300
 DB 939 CCAAGACAGTGTGACTGCTGAGGACATGTGACAGCTGGAGAGTGGGACAGGGTCC 998
 QY 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
 DB 999 ATGATGATGGCGGTGAGCTTATATCATGGAAGACCTCTCACTTAATTAAGATCTT 1058
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnGly 340
 DB 1059 GGGCTGCGTCCAAAGAGACTCTCGCGGTGTCTGAGTCACTGAGCAAGAAAGTGGCA 1118
 QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
 DB 1119 GTTGTGCTTCCAGTATTAATCACTTACCAAGGTAATATTTCCAACTACAGCTGTG 1178
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
 DB 1179 ATGAGTGTGACGAGGAACCTTCTTACCACTGGGCTGCATTCACAGGAGAAAG 1238
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 DB 1239 GCCAGGCGCATCAGAGAGAGTATAGCTGTGAGCCCTCCAAATATACACAGGTC 1298
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnThrProIleGlnAlaGlyValProGlyVala 420
 DB 1299 CTGAGCCATGAGAGAGGACAGACATCACTTTGAGTCAAGCTGAGTGCCTGGAGCC 1358
 QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisIleSerHisGlyAspThrMet 440
 DB 1359 AGTCTACTGATGACTTATCAAGATTAATTTCTTCTTCATCACTCCACAGGAGACCACTG 1418
 QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaAlaValIleProAlaValIleSerTyr 460
 DB 1419 ACTGTCATGATCCAAAGACAGATGAATGTCTGTGCTGTGGCTGTGTTCTTAT 1478
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
 DB 1479 GTTGTGCAAGCATGAGAAATGCTGCTAGTCC 1514

RESULT 2
 AB092016
 ID AB092016 standard; cDNA; 1851 BP.
 XX
 AC AB092016;
 XX
 DT 04-OCT-2002 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 13.
 XX
 KM Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic;
 KM antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
 KM neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
 KM antidiabetic; antidiabetic; antidiabetic; antidiabetic;
 KM immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KM cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KM stem cell; growth factor; nervous system disease; neuropathy;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KM osteoporosis; severe combined immunodeficiency; SCID; infection;
 KM multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002065394-A1.
 XX
 PD 30-MAY-2002.
 XX
 PF 22-DEC-2000; 2000US-0745763.
 XX
 PR 18-MAR-1998; 98US-0040963.
 XX
 PA (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R. A.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (SPAUD/) SPAUDING V.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 XX
 DR WPI: 2002-582343/62.
 DR P-PSDB: ABP61800.
 XX
 PT Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis
 PT
 XX
 PS Claim 50; Page 113-114; 284pp; English.
 XX
 CC The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (cDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (I) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, antinflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,

CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polynucleotide of the invention.
 XX
 SO Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 2,096-236 Length: 1851
 Score: 2417.00 Matches: 472
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 24
 US-09-745-763-36 (1-472) x AB092016 (1-1851)
 QY 1 MetLysPheLeuIlePheAlaPheGlyValHisLeuSerLeuCySerGly 20
 DB 99 ATGAAATTCCTTATCTTCGCACTTTTCGGTGGTCTTACCTTTATCCTGTGCTGGG 158
 QY 21 LysAlaIleCySerLysAengIylSerLysArgThrPheGluGluIle 40
 DB 159 AAACCTATATGCAAGATGCGATCTCTAAGAGCACTTTGGAAGATTAAGAAATA 218
 QY 41 AlaSerCyGlyAapValAlaValAlaIleLeuLeuValAlaValAlaValAlaVal 60
 DB 219 GCCAGCTGTGAGATGTTGCTTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCGAG 278
 QY 61 AsnArgSerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80
 DB 279 AACGATCTCTATGAGCATGCGATCTCTGCTGATATCTGTGAGCCAGCTAGTGGC 338
 QY 81 SerLysAnLeuGluLysAlaIleGlnIleMetTyrGlnAnLeuGlnAnSpGlyLeu 100
 DB 339 TCCAAAGAACTTAAAGAAAGCATCAATATATGTAACCAAACTCGCAGCAAGATGGGCTG 398
 QY 101 GluLysValHisLeuGluProValArgIleProHisIleProGluArgGlyGluGluSerAla 120
 DB 399 GAGAAAGTTCACTCGGAGCCAGTGAATATCCCATCGGAGGAGGAGAAATCAAGCT 458
 QY 121 ValMetLeuGluProArgGlyHisLeuValAlaIleLeuGlyLeuGlySerSerIleGly 140
 DB 459 GTGATGCTGAGGCCAAGATTCATTAAGATTCCTCTGGGCTTGGCAGCAGCATTTGGG 518
 QY 141 ThrProProGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160
 DB 519 ACTCCTCAAGAAAGCATTAAGCAGAAAGTCTGTGGTGAGCTCTTCGATGAACCTGAG 578
 QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValIleYrAnGlnProTyrIleAsnTyr 180
 DB 579 AGAAGGGGCTCAGAGGAGGAGGAGGAGATTTGTTTATTAACCAACTTATCATCACTAC 638
 QY 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaValAlaValAlaVal 200
 DB 639 TCAAGAGGCTGCAATCCAGAACCAAGGGGCGGTGGAGCTGCCAAGGTGGGGCTTTG 698
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleGln 220
 DB 699 GCATCTCTCATTCATCGATCGGTGGCTCTTCATCTACATGCTCTCACACAGGATTCAG 758
 QY 221 GluTyrGlnAapGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 DB 759 GAATACCAAGATGCGGCGCCAAATTCACACGCTGTATAGGTGAGGAGATGCGAGAA 818
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260


```

Db      819  ATGATGTCAGAAATGCTTCTCATGAGATCAAAATATGTCATTCACACTAAAGTGGGGCA 878
Qy      261  LysThrTyrProAspThrAspSerPheLeuThrValAlaGluIleThiGlySerIleTyr 280
Db      879  AAGACTTACCCAGATACATGATTCCTTCAACATCTAGACAGATATCAGGGGAGCAAAAT 938
Qy      281  ProGluGlnValIleuValSerGlyHisLeuAspSerTTPAspValGlyGlnIleVal 300
Db      939  CCAAGACAGCTTGTACTGCTGATGATCTGACATCTGACAGCTGGAGTGGGAGGATGCC 998-
Qy      301  MetAspAspGlyGlyGlyAlaPheIleSerTpgIuaIleuSerLeuIleLysAspLeu 320
Db      999  ATGATGATGGGGGAGAGCTTTATATCATGAGAGCACTCTCACTTATTAAGATCTT 1058
Qy      321  GlyLeuAArgProLysArgThrLeuArgLeuValLeuTTPThAlaGluGluGlnIleGly 340
Db      1059  GGGCTGCTCCAAAGAGACTCTGCGCTGCTCTGAGATGCGAAGAACAGATGGA 1118
Qy      341  ValGlyAlaPheGlnTyrTyrGlnLeuHisIleValAsnIleSerAsnTyrSerLeuVal 360
Db      1119  GTTGATGCTTCCATGATATATATGATTAACAGAGTAATATTTCCAACTACAGCTTGCTG 1178
Qy      361  MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380
Db      1179  ATGAGTCTGACGAGAACCTTCTTACCCACTGGGGCTGCAATTCACCTGGGAGTGA 1238
Qy      381  AlaArgAlaIleMerGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1239  GCCAGGGCCATCATGAGAGAGGTTTGAAGCCCTGCGAGCCCTCAATATCATCTAGCTC 1298
Qy      401  LeuSerHisGlyGlyGlyThrAspIleAsnPheTTPileGlnAlaGlyValProGlyVal 420
Db      1299  CTGACCCATGAGAGAGGACAGACATCACTTTTGATTCAGACTGGAGTGGCTGGAGCC 1358
Qy      421  SerLeuLeuAspAspLeuTyrIleTyrPhePhePheHisSerHisGlyAspThrMet 440
Db      1359  AGTCACTTGAATGATTAACAGATATTTCTTCCATCACTCCACAGAGACACCAAG 1418
Qy      441  ThrValMetAspProLysGluMetCysValAlaAlaIleValTTPAlaValIleSerTyr 460
Db      1419  ACTGCATGAGATCCAAAGCAGATGAATGTGCTGCTGTTGGCTGTGTGTTCTTAT 1478
Qy      461  ValValAlaAspMetGluGluMetLeuProArgSer 472
Db      1479  GTTGTTGACGACATGAAAGAAATCTGCTAGGTCC 1514

RESULT 3
AAZ98034
ID      AAZ98034 standard; cDNA; 1863 BP.
AC      AAZ98034;
XX      09-MAY-2000 (first entry)
DT      09-MAY-2000 (first entry)
XX      Human secreted protein encoding nucleotide sequence SEQ ID NO:28.
DE      Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX      antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;
XX      tumour; neurodegenerative disorder; developmental abnormality; allergy;
XX      foetal deficiency; blood disorder; immune system disorder; arthritis;
XX      autoimmune disease; hepatic disease; renal disease; inflammation;
XX      Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
XX      infection; AIDS; spinal cord injury; transplant rejection; diabetes;
XX      asthma; sepsis; acne; psoriasis; cardiovascular disorder;
XX      reproductive disorder; gastrointestinal disorder; respiratory disorder;
XX      metabolic disorder; food additive; preservative; ss.
OS      Homo sapiens.
XX      WO200004140-A1.
XX      27-JAN-2000.

```

```

XX      14-JUL-1999; 99WO-US1849.
PF      15-JUL-1998; 98US-0092921.
XX      15-JUL-1998; 98US-0092922.
PR      15-JUL-1998; 98US-0092956.
XX      15-JUL-1998; 98US-0092956.
XX      (HUMA-) HUMAN GENOME SCI INC.
XX      Ruben SM, Komatsu S, G, Duan RD, Rosen CA, Moore PA, Shi Y;
XX      Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;
XX      Mueneski M, Endress GA, Soppet DR;
XX      WPI; 2000-161126/14.
DR      P-PSDB; AA87081.
XX      New isolated human genes, useful for diagnosis and treatment of, e.g.
XX      cancers, neurological or blood disorders
XX      Claim 1; Page 319; 494pp; English.
XX      The polynucleotide sequences given in AAZ98017 to AAZ98108 encode the
XX      human secreted proteins given in AA87064 to AA87223. Human secreted
XX      protein can have activities based on the tissues and cells the genes are
XX      expressed in. Examples of activities include: cytostatic;
XX      immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
XX      antiallergic. The polynucleotides and their corresponding secreted
XX      polypeptides are useful for preventing, treating or ameliorating medical
XX      conditions, e.g. by protein or gene therapy. Also pathological conditions
XX      can be diagnosed by determining the amount of the new polypeptides in a
XX      sample or by determining the presence of mutations in the new
XX      polynucleotides. Human secreted protein s and their polynucleotides can
XX      be used for developing products for the diagnosis or treatment of cancer,
XX      tumours, neurodegenerative disorders, developmental abnormalities and
XX      foetal deficiencies, blood disorders, diseases of the immune system,
XX      autoimmune diseases, hepatic and renal disease, inflammation,
XX      allergies, Alzheimer's disease, behavioural disorder, schizophrenia,
XX      osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX      transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
XX      cardiovascular disorders, reproductive disorders, gastrointestinal
XX      disorders, respiratory disorders and metabolic disorders. The
XX      proteins or polynucleotides can also be used as food additives or
XX      preservatives. The proteins are also useful for identifying their
XX      binding partners. AAZ98008 to AAZ98016 and AA87063 are sequence used in
XX      the exemplification of the present invention.
XX      Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
SQ      Alignment Scores:
XX      Pred. No.: 2,11e-236 Length: 1863
XX      Score: 2417.00 Matches: 472
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 21 Gaps: 0
XX      US-09-745-763-36 (1-472) x AAZ98034 (1-1863)
Qy      1 MetLysPheLeuIlePheAlaPheGlyGlyValHisLeuSerLeuCysSerGly 20
Db      99 ATGAAATTCCTTATCTTCGATCTTTCGCGGTGTTTCACTTTATCTCTGCTGGG 158
Qy      21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40
Db      159 AAACCTTATGCAAGATGAGCTCTTAAGAGGACTTTTGAAGAAATTAAGAGAAATA 218
Qy      41 AlAserCysGlyAspValAlaLysAlaIleIleLeuAsnAlaValTyrGlyLysAlaGln 60
Db      219 GCCAGCTGTGAGAGATTTGCTTAAGCAATCATCACTAGCTGTTTATGTAAGCCAG 278
Qy      61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db      279 AACAGATCTATAGGAGCATTTGGCACTTCTGGTTGATATCTGTGGACCAAGCTGAGTGGC 338

```

QY 81 SerLyAsnLeuGluValAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
 DB 339 TCAGAAAGCCTTAGAAAAAGCCATCCAAATTATGTATCCAAAACCTGCAGACAGATGGGCTG 398
 QY 101 GtulySValHsLeuGluProValArgIleProHisTTPGluArgGlyGluGluSerAla 120
 DB 399 GAGAAAGTTCACCTGGACCCAGTGAATACCCCACTGGGAGAGGGGAGAAATACAGCT 458
 QY 121 ValMetLeuGluProArgIleHisIleValIleLeuGlyLeuGlySerSerIleGly 140
 DB 459 GTGATGCTGGAGCCAGAAATTCATAGATAGCATCTGGGTCTTGAGAGACATGGG 518
 QY 141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
 DB 519 ACTCTCCAGAAAGCATTACAGCAAGTCTGGGTGGAGACTCTTCGATGAACTGAG 578
 QY 161 ArgArgAlaSerGluAlaArgGlyValIleValValIleValProTyrIleAsnTyr 180
 DB 579 AGAAGGGCTCAGAAAGCAGAGAGAGATGTGTTTATACCAACCTTACATCACTAC 638
 QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaIleValValIleVal 200
 DB 639 TCAAGAGCGGTGCAATACCAAGCAGGGGCGGTGAGAGCTGCCAAGGTTGGGCTTGG 698
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
 DB 699 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCTCACACAGGATTTG 758
 QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 DB 759 GAATACCAAGATGGCGTCCCAAGATTCACACAGCTGTATTCAGGTGAGAGATGCAGAA 818
 QY 241 MetMetSerArgMetAlaSerHisGlyIleIleValIleGlnLeuLysMetGlyVal 260
 DB 819 ATGATGTCAGAAAGTCTTCATGAGATCAAAATGTCATTCAGTCAAGATGGGGCA 878
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280
 DB 879 AAGACCTACCAAGTACTGATCTCTTCAACACTGTAGACAGATCATCGGAGCAAAAT 938
 QY 281 ProGluGlnValIleLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnIleVal 300
 DB 939 CCAGAACAGGTTGTAAGTGTAGTACATCTGCAGACTGCAGATGTTGGCAGAGGTGCC 998
 QY 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
 DB 999 ATGATATATGGCGGTGAGCCTTATATATCATGGAGACACTCTCATATTAAAGATCTT 1058
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThrAlaGluGluGlnIleGly 340
 DB 1059 GGGCTGCTCCAAAGAGAGACTCTGGGCTGGTCTCTGACTGCAGAAAGCAAGGTGGA 1118
 QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisIleValAsnIleSerAsnTyrSerLeuVal 360
 DB 1119 GTTGCTGCTCCAGTATTATCAGTTACACAAAGTAATATTTCACATACAGCTGGTG 1178
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
 DB 1179 ATGAGGCTGACGAGCAACCTTCTTACCACTGGGCTGCATTCACGAGCAAGAAAG 1238
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuGlnIleProLysAsnIleThrGlnVal 400
 DB 1239 GCCAGGGCCATCATGAGAGAGGATTATGAGCCTGCTGCAGCCCTCAATATACATCAGGTC 1298
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrPileGlnAlaGlyValProGlyVal 420
 DB 1299 CTGAGCCATGAGAGAGGACAGACATCAACTTTGGATCCAAAGCTGGAGTCTGGAGCC 1358
 QY 421 SerLeuLeuAspAspLeuTyrIleTyrPhePhePheHisIleSerHisGlyAspThrMet 440
 DB 1359 AGCTACTTATGATGACTTATACAAAGTATTTCTTCCATCACTCCACGAGAGACACCATG 1418

QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValThrAlaValValSerTyr 460
 DB 1419 ACTGTCATGATCCAAAGACAGATGATGCTGCTGCTGTTGGCTGTTGTTCTTAT 1478
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
 DB 1479 GTTGTTCCAGCATGAGAAAGAAATGCTGCTTAGGTC 1514
 RESULT 4
 AAD11647 ID AAD11647 standard; cDNA; 1863 BP.
 AC AAD11647;
 DT 24-SEP-2001 (first entry)
 XX
 DE Human secreted protein-encoding gene 18 cDNA clone HRACT35, SEQ ID NO:28.
 KW Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KW cell culture; chemotaxis; vulnerability; binding partner identification;
 KW gene therapy; ss.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 99..1517
 FT /*tag= a
 FT /product= "Human secreted protein precursor"
 FT sig_peptide 99..170
 FT /*tag= b
 FT mat_peptide 171..1514
 FT /*tag= c
 FT /product= "Mature human secreted protein"
 FT
 XX MO200151504-A1.
 XX
 PD 19-JUL-2001.
 XX
 PF 12-JAN-2001; 2001WO-US00911.
 XX
 PR 13-JAN-2000; 2000US-0482273.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 FI Ruben SM, Komatsuolis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
 FI Laflair DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
 PI Andrews GA, Mucenski M, Edner R;
 XX
 DR WPI: 2001-425865/45.
 XX P-PSDB; AAB06058.
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 XX
 PS Claim 1; Page 686; 864pp; English.
 CC AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
 CC protein genes, and AAB06041-AAB06132 represent the proteins they encode.
 CC AAB06133-AAB06205 represent human secreted protein fragments.
 CC The secreted proteins and their genes are useful for preventing, treating
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 CC Pathological conditions can be diagnosed by determining the amount of the
 CC new protein in a sample or by determining the presence of mutations in
 CC the new genes. Specific uses are described for each of the 71 genes,
 CC based on the tissues in which they are most highly expressed, and include
 CC developing products for the diagnosis or treatment of proliferative

disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiosenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
 CC The present sequence represents a human secreted protein-encoding cDNA of the invention.

XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Alignment Scores:

Pred. No.: 2,116-236 Length: 1863
 Score: 2417.00 Matches: 472
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-745-763-36 (1-472) x AAD11647 (1-1863)

QY 1 MetLysPheLeuIlePhePheLeuIleValHisLeuLeuSerLeuCySerGly 20
 DB 99 ATGAATTCCTTATCTGCAATTTTCGGTGATGTCACCTTTATCCCTGCTCTG 158
 QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluLurle 40
 DB 159 AAAGCTATATGCAAAATGCGCATCTCTAAGAGCATTTTGAAGAAATTAAGAAATA 218
 QY 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
 DB 219 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTGATGTTATGTTAAACCCAG 278
 QY 61 AsnArgSerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80
 DB 279 AACGATCTTATGACGAGATGGCACTTCTGTTGATCTGTGACCCAGACTGAGTGC 338
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100
 DB 339 TCAGAAACCTTGAAGAAAGCCATCCAAATTATGTCACAAACCTTCAGCAAGATCGGCTG 398
 QY 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerAla 120
 DB 399 GAGAAAGTTTCACTGAGCCAGTGAATATCCCACTGGGAGAGGGAGAAATCAGCT 458
 QY 121 ValMetLeuGluProArgIleHisLysValIleAlaIleLeuGlyLeuGlySerSerIleGly 140
 DB 455 GTGATGCTGAGACCAAGATTCATTAAGATAGCCATCTCGGTCTTGAGACGACCTTGG 518
 QY 141 ThrProGluGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160
 DB 519 ACTCTCCAGAAAGGATTCACGACAGATGTCGTGGTGGACCTCTTTCAGATGAATCGAG 578
 QY 161 ArgAlaGlaSerGluAlaArgGlyLysIleValAlaTyrAsnGlnProTyrIleAsnTyr 180
 DB 579 AGAAGGCGCTCAGAGCAAGAGGAGAAATGTTGTTTATACCAACCTTACATCAATAC 638
 QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValAlaIleAlaLysValGlyAlaLeu 200
 DB 639 TCAGAGAGGTGCATATCCGAACGAGGGGGGCGTGAAGCTGCCAAGGTGGGGCTTTTG 698
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220

DB 699 GCATCTCTCATTCGATCGGTGGCTCTCTTCATCTCAAGTCTCCACAGAGTATTCAG 758
 QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 DB 759 GAATACCAAGATGGCGTGGCCCAAGATTCACACCTGTATATACGGTGGAAAGATGCAGAA 818
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 DB 819 ATGATGTCAGAAATGGCTTCTCATGGGATCCAAATTTGATTCATTCAGTAAAGATGGGCA 878
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
 DB 879 AAGACTTACCCAGATATGATGATCTTCTCAACACTGATACAGAGATCATGGGACCAATAT 938
 QY 281 ProGluGlnValIleLysValSerGlyHisLeuAspSerThrPaspValGlyGluAla 300
 DB 939 CCAGAACAGTTTGAATCGTGTAGTGACATCTGACAGCTGGGATGTTGGGCAAGGTTGCC 998
 QY 301 MetAspAspGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
 DB 999 ATGATGATGGCGGTGAGACCTTTATATCATGGAGACATCTCATTTATTAAGATCTT 1058
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnGlyGly 340
 DB 1059 GGGCTGCGTCCAAAGAGAGACTCTGGCTGGTCTTGAGACTGCAGAAACAGAGTGA 1118
 QY 341 ValGluLysPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
 DB 1119 GTTGTGCTTCCTCCAGTATTCACATTCACAGGATTAATTTCCAACTGACAGTGGTGG 1178
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
 DB 1179 ATGAGTCTGACGAGAACTTTCTTACCACTGGCTGCATTCATCGGAGAGTGAAG 1238
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 DB 1239 GCCAGGCGCATCATGAGAGAGATGATGAGCTGTCGACCCCTCAATATCATCTCAGGTC 1298
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420
 DB 1299 CTGAGCCATGAGAGAGGAGACATCATCTTGTGATCCAAAGTGGAGTGCCTGGAGCC 1358
 QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
 DB 1359 AGTCTACTGATGATCTTATACAGATATTTCTTCTTCATCACTCCACGAGACCATG 1418
 QY 441 ThrValMetAspProLysGlnMetAspValAlaAlaIleValTPAlaValIleSerTyr 460
 DB 1419 ACTGTCAATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTCTTAT 1478
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
 DB 1479 GTTGTTCACATGACATGAAGAAATGCTGCTAAGTCC 1514
 DB 1514

RESULT 5
 ABK69743
 ID ABK69743 standard; cDNA; 1863 BP.
 AC ABK69743;
 DT 15-JUL-2002 (first entry)
 XX Human secreted protein gene 18 #1.
 XX Human; ss; gene; secreted protein; gene therapy; immunosuppressive;
 KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiac;
 KW vasotrophic; cerebroprotective; nootropic; neuroprotective; antibacterial;
 KW vitruicide; fungicide; ophthalmological; autoimmune disease; neoplasm;
 KW rheumatoid arthritis; hyperproliferative disorder; cerebral arrest;
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
 KW ocular disorder; corneal infection; wound healing; skin aging;

KW epithelial cell proliferation; food additive.
 XX Homo sapiens.
 XX MO200226931-A2.
 XX PD 04-APR-2002.
 XX PF 24-SEP-2001; 2001WO-US29871.
 XX PR 25-SEP-2000; 2000US-234925P.
 XX PR 12-JAN-2001; 2001WO-US00911.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsuji S, Duan DR, Rosen CA, Moore PA, Shi Y,
 PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR,
 PI Endress GA, Muceneki W, Ebner R,
 XX WPI; 2002-362489/39.
 DR P-PSDB; ABG33880.
 XX
 PT Novel 71 isolated secreted polypeptides and polymucleotides encoding
 PT the polypeptides, useful for treating Huntington's disease, sepsis,
 PT meningitis, thrombocytopaenia, haemolytic anaemia, rheumatoid arthritis,
 PT asthma
 XX
 PS Claim 1; Page 1170; 1478pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (or its
 CC fragment, homologue complement or allelic variant) encoding a human
 CC secreted protein (and its fragment, domain, epitope, variant, secreted
 CC form and species variant). Also included are a recombinant vector
 CC comprising the nucleic acid, a recombinant host cell comprising the
 CC vector, an antibody against the secreted protein, a recombinant host cell
 CC that expresses the secreted protein and a method of identifying a binding
 CC partner of the secreted protein. The nucleic acid and protein are used to
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
 CC for example autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
 CC ocular disorders e.g. corneal infection. Many other diseases and
 CC disorders are listed in the specification. The polypeptides can also be
 CC used to aid wound healing in an epithelial cell proliferation, to prevent
 CC skin aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities. The present
 CC sequence encodes a novel human secreted protein of the invention.
 XX
 SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 2,11e-236 Length: 1863
 Score: 2417.00 Matches: 472
 Percent Similarity: 100.00% Conservatave: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-745-763-36 (1-472) x ABK69743 (1-1863)
 QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20
 DB 99 ATGAATTCCTTATCTTGCAATTTTCGGTGGTGTACCTTTATCCCTGGCTGGG 158
 QY 21 LysAlaIleCyGlyAsnGlyYIleSerYsArgThrPheGluGluIleLysGluGluIle 40
 DB 159 AAAGCTATATGCAAGAAATGGCATCTCTAAGAGCACTTTGAAGAAATAAAGAAATA 218

QY 41 Alasercysglaspvalalalyalallelelsenuvalavaltyrglylsalagln 60
 DB 219 GCCAGCTGTGGAGATGTGTGCTTAAAGCATCATCACTACTGTTATGTTAAAGCCAG 278
 QY 61 AsnArgSerTyrgluargleuvalaleuvalaspThrValGlyProargLeuSerGly 80
 DB 279 AACGATCTATGAGCAGATTGGCACTTCTGTGATGATCTGTGGACCAACTGAGTGGC 338
 QY 81 SerLysAsnLeuGlyLysAlaIleGlnIleMetTyrglnAsnLeuGlnIleAspGlyLeu 100
 DB 339 TCCAGAACCTTAGAAAAAGCCATCCAAATTATGTAACAAAACCTGCAGCAAGATGGCCTG 398
 QY 101 GluLysValHisLeuGluProvalArgIleProHisTrpGluArgGlyGluGluSerAla 120
 DB 399 GAGAAATTCACCTGGAGCCAGTGAATATCCCACTGGAGAGGGAGAAATTCAGCT 458
 QY 121 ValMetLeuGluProargIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
 DB 459 GTGATCTGGAGCCAAATTCATTAAGATAGCCATCTGGGCTTGGCAGACGATGGG 518
 QY 141 ThrProGluGluIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
 DB 519 ACTCTCCAGAAAGCATTTACAGCAGAACTTGTGTGTGACCTTTTCGATGAATGTCAG 578
 QY 161 ArgArgLaseGluAlaArgGlyLysIleValValTyraGlnProTyrlleAsnTyx 180
 DB 579 AGAAGGCTCCAGAGCAAGAGAGGAGATGTTGTTTATACCACTTACATCACTAC 638
 QY 181 SerArgThrValGlnTyraGlnArgGlnIleValAlaGluAlaLysValGlyAlaLeu 200
 DB 639 TCAAGAGCGGTGCATATCCCAACGCGAGGGCGGTGGAACCTGCAGAGTTGGGCTTTG 698
 QY 201 AlasercleuIleAgsSerValAlaSerPheSerIleTyserProHisTrpGlyIleGln 220
 DB 699 GCATCTTCATTCGATTCGATTCGGCTCTCTCTTCATTCACAGTCTCCACAGGATTCAG 758
 QY 221 GluTyrglnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 DB 759 GAATACCGATGGCGTCCCAAGATTCCAACGCTGTATTAGGTGGAGAGATGCAGAA 818
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 DB 819 ATGATGTCAGAAATAGGCTTTCATGAGGATCAAAATTCATTCAGTCAAGATGGGGCA 878
 QY 261 LysThrTyProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyx 280
 DB 879 AAGACCTACCCAGATGATGATCTCTTCACACTGTACAGAGATCACTGGAGAGCAATAT 938
 QY 281 ProGluGlnValValleuValSerGlyHisLeuAspSerTrpAspValGlyGlnIleVala 300
 DB 939 CCAGAACAGGTTGATGTCAGTGCAGTGCAGTGCAGTGGAGTGGGACAGGAGGCC 998
 QY 301 MetAspAspGlyGlyValAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
 DB 999 ATGATGATGGCGGTGAGCCCTTATATCATGGGAACACCTCATTTAAAGATCTT 1058
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnIleGly 340
 DB 1059 GGGCTGTGGTCCAAAGAGACTCTGGCTGGTGTCTGAGACTGCAGAAAGAAAGAGTGGGA 1118
 QY 341 ValGlyAlaPheGlnIleTyrglnLeuHisLysValAsnIleSerAsnTyxSerLeuVal 360
 DB 1119 GTTGATGCTTCCAGATATATCATATACATGACAAAGGTAAATTTTCAACTACAGTGGTG 1178
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
 DB 1179 ATGAGCTGAGCGGAGAACTTCTTACCACTGGGCTGCAGATTCATCGGAGAGAAAG 1238
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 DB 1239 GCCAGGGCCATCAAGAGGAGGATATAGCCTGTGAGGCCCTCCAAATATCACCTCAGTGC 1298
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnIleGlyValProGlyAla 420

Db 1299 CTGAGCCATGAGAGGAGACATCACTTTGGATCCAAAGCTGAGTGCCTGAGACC 1358
 Qy 421 SerLeuLeuAAspAAspLeuTyrTyrPhePhePheHisSerHisGlyAspPheMet 440
 Db 1359 AGTCTACTTGAGATGATTAACAGATATTTCTTCCATCACTCCACGAGACCCATG 1418
 Qy 441 ThrValMetAAspProGlyGlnMetAsnValAlaAlaValTyrAlaValSerTyr 460
 Db 1419 ACTGTCATGATCCAAAGACAGATGATGTGCTGCTGCTTTGGGCTGTCTTCTTAT 1478
 Qy 461 ValValAlaAspMetGlnGlnMetLeuProArgSer 472
 Db 1479 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1514
 RESULT 6
 ACC50817
 ID ACC50817 standard; cDNA; 1863 BP.
 AC ACC50817;
 XX
 DT 12-JUN-2003 (first entry)
 DE Human secreted protein coding sequence, SEQ ID 484.
 KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;
 KW vulnereary; antiinflammatory; nootropic; neuroprotective;
 KW antiparkinsonian; gene therapy; human; cardiovascular disorder;
 KW gene; ss.
 XX
 OS Homo sapiens.
 PM MO200295010-A2.
 PD 28-NOV-2002.
 PF 19-MAR-2002; 2002WO-US09785.
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM;
 DR WPI; 2003-129429/12.
 PT Novel human secreted proteins, useful for detecting, preventing,
 PT diagnosing, prognosticating, treating and/or ameliorating
 PT cardiovascular disorders such as arrhythmia -
 XX
 PA Claim 21; SEQ ID 484, 1861bp; English.
 XX
 XX The present invention relates to novel human secreted proteins
 CC (AB474733-AB48145) and their coding sequences (ACC50344-ACC50856). The
 CC proteins and their coding sequences are useful for the preparation of a
 CC diagnostic or pharmaceutical composition for diagnosing or treating a
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
 CC coronary arteriosclerosis and myocardial ischemia), neural disorders,
 CC immune system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of haematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of

CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism.
 CC Note: The sequence data for this patent was published in electronic
 CC format and is available from WIP0 at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,11e-236 Length: 1863
 Score: 2417.00 Matches: 472
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-745-763-36 (1-472) x ACC50817 (1-1863)
 Qy 1 MetLysPheLeuLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20
 Db 99 ATGAATTCCTTATCTTGGATTTTGGTGGTGTTCACCTTTATCCTGCTGCTGG 158
 Qy 21 LysAlaIleCySerLysAsnGlyIleSerLysArgThrPheGlnGlnIleLysGlnIle 40
 Db 159 AAACCTATATGCAAGATGGCATCTCTAAGAGGACCTTTTGAGAAATTAAGAAATA 218
 Qy 41 AlaserCySerGlyAspValAlaValAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
 Db 219 GCCAGCTGTGAGATGTTGCTTAAGACATCACTCAAGCTGTTTATGTTAAGCCCG 278
 Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
 Db 279 AACAGATCTTATGAGCGCATTTGGCACTTCTGTTATATCTGTGAGCCAGCATGAGTGC 338
 Qy 81 SerLysAsnLeuGlnLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100
 Db 339 TCCAAAGAACCTAAGAAAAAGCCATCCAAATTAATGACAAACCTGACGCAAGATGGCTG 398
 Qy 101 GlnLysValHisLeuGlnProValArgIleProHisTrrProGluArgGlyGlnSerAla 120
 Db 399 GAGAAAGTTCACCTGAGGCCGTGAGATACCCACCTGGAGAGGAGAGAAATTCACCT 458
 Qy 121 ValMetLeuGlnProArgIleHisIleValIleAlaIleLeuGlySerSerIleGly 140
 Db 459 GTGATGCTGAGGCCAAGATTCATTAAGATGACCATCTCGGCTTGGAGCAGCATTTGGG 518
 Qy 141 ThrProProGlnGlyIleThrAlaGlnValLeuValValThrSerPheAspGlnLeuGln 160
 Db 519 ACTCCTCCAGAGGACATTACAGCAAGTTCTGTGTGTGACCTCTTGCATGAACCTCGAG 578
 Qy 161 ArgArgAlaSerGlnAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
 Db 579 AGAAGGGCTTCAGAAAGCAAGAGAGAGATTTGTTTATTAACCACTTACATCAACTAC 638
 Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGlnAlaAlaLysValGlyAlaLeu 200
 Db 639 TCAAGAGCGTGAATCCAGAACCAAGCGAGGGCGGTGAGACCTGCAGAGTTGGGCTTTG 698
 Qy 201 AlaserLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleThrGlyIleGln 220
 Db 699 GCATCTCTCATTTGATTCATGCTGTGGCTCTTCCATCTCAACATCTCTCACACAGGATTCAG 758
 Qy 221 GlnLysArgAspGlyValProLysIleProThrAlaCysIleThrValGlnAspAlaGln 240
 Db 759 GAATTCACAGATGCGTCCCAAGATTCACACCTGTATTAAGGTGAGAAATGACAGAA 818
 Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 Db 819 ATGATGTCAGAAATGGCTTCATGAGGATCAAAATGTGCATTACGCTAAAGATGGGGCA 878
 Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGlnIleThrGlySerLysTyr 280
 Db 879 AAGACCTTACAGATGATGATTCCTTCAACACTGTAGCAGAGATCAGTGGAGCAAAATAT 938

QY 281 ProgluValValIleuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
 DB 999 CCAAGACAGGTTGTAAGTGTGATGAGATCTGGACACTGGAGCTGGAGCTGGAGGAGGTC 998
 QY 301 MetAspAspGlyGlyValAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
 DB 999 ATGGATGATGGCGGTGAGCCCTTATATCATGGAGACCTCTCAATTATAAAGATCTT 1058
 QY 321 GlyLeuAspProLysArgTrpLeuArgLeuValLeuTrpThrAlaGluGlnGlyGly 340
 DB 1059 GGGGCTGCTCCAAAGAGACTCTGGGCTGGCTGTGACTGACGCAAGAAACAGTGTGA 1118
 QY 341 ValGlyAlaPheGlnTrpGlnLeuHisLysValAsnIleSerAsnTrpSerLeuVal 360
 DB 1119 GTTGATGCTTCCAGATATATCATGATACAGAAAGTAATTTCCAACTACAGCTGGATG 1178
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
 DB 1179 ATGGAGTCTGACGAGAACCTTCTTACCCTGGGCTGCAATTCACTGGCAGTGAAGA 1238
 QY 381 AlaArgAlaIleMetGluGlnValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 DB 1239 GCCAGGGCCATCATGAGAGAGGTTATGAGCCCTGCGACGCCCTCAATATCATCTCAGGTC 1298
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
 DB 1299 CTGACCCATGAGAGAGAGAGACATCACTTTGGATCCAGCTGGAGTGGCTGGAGACC 1358
 QY 421 SerLeuLeuAspAspLeuTrpLysThrPhePhePheHisHisSerHisGlyAspThrMet 440
 DB 1359 AGTCACTTGAATGACTTATACAGATATTTCTTCCATCACTCCACAGAGACCACTG 1418
 QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaAlaValTPAlaValAspSerTrp 460
 DB 1419 ACTGCATGATGCCAAAGCAGATGATGTGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1478
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
 DB 1479 GTTGTTCAGACATGAGAAAGAAATGCTGCTTACGTTCC 1514

RESULT 7
 ABZ71453
 ID ABZ71453 standard; cDNA, 1863 BP.

AC ABZ71453;

DT 04-APR-2003 (first entry)

DE Secreted protein-encoding gene 142 cDNA clone HRACU35, SEQ ID NO:274.

Human; secreted protein; digestive disorder; gastrointestinal disorder;
 mouth; oesophagus; stomach; small intestine; large intestine; liver;
 biliary tract; pancreas; cancer; tumor; hyperproliferative disorder;
 immune disorder; inflammation; infection; wound healing; drug screening;
 chromosome identification; chromosome mapping; cytostatic; gene therapy;
 antiinflammatory; immunosuppressive; veterinary; chromosome 8q22.2;
 gene; ss.

OS Homo sapiens.

PN WO200276488-A1.

PD 03-OCT-2002.

PE 19-MAR-2002; 2002WO-US08276.

PR 21-MAR-2001; 2001US-277340P.

PR 19-JUL-2001; 2001US-306171P.

PR 13-NOV-2001; 2001US-331287P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;
 XX WPI: 2003-029900/02.
 DR P-PSDB; ABR00274.
 XX
 PT New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 PT e.g. gastrointestinal diseases and disorders, or cancers -
 XX
 PS Claim 21; Page 918, 1216pp; English.

CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein-
 CC encoding cDNA clone of the invention.

SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other:

Alignment Scores:

Pred. NO.: 2,11e-236 Length: 1863
 Score: 2417.00 Matches: 472
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-745-763-36 (1-472) x ABZ71453 (1-1863)

QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCysSerGly 20
 DB 99 ATGAAATTCCTTATCTTCGCAATTTTGGTGGTTCACCTTTATCTCTGCTGGG 158
 QY 21 LysAlaIleCysLysAsnGlyLysSerLysArgThrPheGluGluIleLysGluIle 40
 DB 159 AAAGCTATATGCAAGAAATGGATCTTAAAGGACCTTTGAAGAAATGAAGAAATA 218
 QY 41 AlaSerCysGlyAspValAlaLysAlaIleLeuLeuAlaValTrpGlyLysAlaGln 60
 DB 219 GCCAGCTGTGAGATGTGCTAAAGCATCATCACTAGCTGTTATGGTAAAGCCAG 278
 QY 61 AsnArgSerTrpGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
 DB 279 AACGATCTTATGAGCATGGCACTTCTGGTTGATATCTGTGAGACCAGACTAGTGGC 338
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTrpGlnAsnLeuGlnAspGlyLeu 100
 DB 339 TCAGAGAACTTAAAGAAAGCCATCAATTTATGCAAAACCTGACAGCAAGATGGCTG 398
 QY 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluGluSerAla 120
 DB 399 GAGAAAGTTACCTGAGAGCCAGTGAATATCCCACTGGAGAGAGGAGAAATGAGCT 458
 QY 121 ValMetLeuGluProArgGlyLeHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
 DB 459 GTGATGCTGAGGCCAAATTCATAGATAGCCATCTCTGGGCTTGGCAGCAGCATTTGG 518
 QY 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160


```

Db      |||
519 ACTCTCCAGAAAGGATTACAGCAGAGTTCTGGTGGTACCTCTTGATCACTGCAG 578
Qy      |||
161 ArgAGAlaSerGluAlaArgGlyLysIleValValIYrAenGlnProTyrIleAenTyr 180
Db      |||
579 AGAAGGGCTCTGAAAGCAGAGAGAAAGATGGTGGTTTAAACCAACTTACCTCAACTAC 638
Qy      |||
181 SerAGTThrValGlyTyrArgThrGlnGlyAlaValAlaIleValValGlyAlaLeu 200
Db      |||
639 TCAGAGCGGTGCATACCGAACGAGGGGGCGGTGGAGCTGCCAGGTTGGGCTTGG 698
Qy      |||
201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db      |||
699 GCATCTCTCATTCGATCCGTGGCTCTCTCTCATCTACAGTCTCAGACAGATATCG 758
Qy      |||
221 GltYrGlnAenGlyValProLysIleProThrAlaCysIleThrValGlnAenAlaGlu 240
Db      |||
759 GAATACCGAGATGGGTGCCAAGATTCACACAGCCGTGTATTCAGGTGAGATGCAGAA 818
Qy      |||
241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db      |||
819 ATGATGTCAAGATGGCTTCTTCATGGATCAAAATTTGATTCAGCTAAAGTGGGGCA 878
Qy      |||
261 LysThrTyrProAspThrAspSerPheAenThrValAlaGluIleThrGlySerLysTyr 280
Db      |||
879 AAGACCTACCCAGATACGATTCCTTCAACACTGAGAGAGATCAGCTGGAGCAAAATAT 938
Qy      |||
281 ProGlnGlnValIleValLeuValSerGlyHisLeuAenSerTTPAspValGlyGlnGlyAla 300
Db      |||
939 CCAGAACAGGTGTGACTGGTCACTGAGACATCTGACAGCTGGAGATGGGCGAGGTGCC 998
Qy      |||
301 MetAspAspGlyGlyGlyValAlaPheIleSerTTPGluAlaLeuSerLeuIleLysAspLeu 320
Db      |||
999 ATGGATGATGGCGGTGGAGCTTTATATCATGGAGACACTTCACTTATTTAAATCTT 1058
Qy      |||
321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTyrPThrAlaGlnGlnGlnGly 340
Db      |||
1059 GGGCTGCGTCCAAAGAGGACTCTGGCGGTGGTGGTCTGGAGCTGAGAAAGCAAGTGGGA 1118
Qy      |||
341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAenIleSerAenTyrSerLeuVal 360
Db      |||
1119 GTTGGTGGCTTCCAGTATTTATAGTTATCAACAGGTAATAATTTCCAACTACAGTCTGGTG 1178
Qy      |||
361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlnLys 380
Db      |||
1179 ATGGAGTCTGACCGCAGAACTTTTACCCACTGGGCTGCAATTCACTGGCAGTAAAG 1238
Qy      |||
381 AlaArgAlaIleMetGlnGlnValMetSerLeuLeuGlnProLeuAenIleThrGlnVal 400
Db      |||
1239 GCCAGGGCCATCATGAGAGAGGTATAGCTGCTGCAGAGCCCTCAATATCACTCAGGTC 1298
Qy      |||
401 LeuSerHisGlyGlnGlyThrAspIleAenPheTTPIleGlnAlaGlyValProGlyAla 420
Db      |||
1299 CTGACCCATGAGGAGGAGCAGACATCACTTTTGATTCACAGCTGGAGTGGCTTGAGACC 1358
Qy      |||
421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
Db      |||
1359 AGTCTACTTGATGATCTTATACAGATATTTCTTCTTCATCACTCCACAGAGACACCATG 1418
Qy      |||
441 ThrValMetAspProLysGlnMetAenValAlaAlaAlaValTTPAlaValIleSerTyr 460
Db      |||
1419 ACTGTCATGATGCCAAAGACAGATGATGTGCTGCTGGCTTTGGGCTGTTTCTTAT 1478
Qy      |||
461 ValValAlaAspMetGlnGlnMetLeuProArgSer 472
Db      |||
1479 GTTGTTCAGACATGAGAAAGAAATGCTGCTAGGTCC 1514

```

RESULT 8

AAZ58313

ID AAZ58313 standard; cDNA; 1884 BP.

XX AC AAZ58313;

XX

```

DT      08-MAY-2000 (first entry)
XX      Human peptidase NALAD-ase IV cDNA.
DE      NALAD-ase IV; N-acetylated alpha-linked acidic dipeptidase; human;
XX      chromosome 8q21.3; prostate cancer; neurodegenerative disease;
XX      Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
XX      peripheral neuropathy; Huntington's disease; acute brain injury;
XX      multiple sclerosis; peripheral nerve trauma; ischemia; dementia;
XX      gene therapy; diagnosis; noctropic; neuroprotective; neuroleptic;
XX      antiparkinsonian; anticonvulsant; vasotropic; ss.
XX      Homo sapiens.
OS      Key      Location/Qualifiers
FH      CDS      149..1567
FT      /*tag= a
FT      MO200004157-A2.
PN      27-JAN-2000.
PD      14-JUL-1999; 99WO-GB02241.
PF      14-JUL-1998; 98GB-0015284.
PR      14-JUL-1998; 98GB-0015284.
PA      (JANNC ) JANSSEN PHARM NV.
PI      Pangalos M, Neefs JEFM, Peeters DCG;
XX      WPI; 2000-182424/16.
DR      P-PSDB; AAY58879.
XX      New human N-acetylated alpha-linked acidic dipeptidases for treating
PT      neural disorders e.g. Alzheimer's disease, schizophrenia and
PT      Parkinson's disease.
XX      Claim 7; Fig 5; 95pp; English.
XX      The present sequence is that of cDNA coding for human
XX      N-acetylated alpha-linked acidic dipeptidase IV (NALAD-ase IV,
XX      see AAY58879). The cDNA was obtained from a gall bladder cDNA
XX      library. Analyses of the open reading frame predicts a type II
XX      integral membrane protein with 5 potential N-glycosylation sites.
XX      The NALAD-ase II gene was mapped to chromosome 8q21.3.
XX      NALAD-ase IV expression was low in all tissues examined by RT-PCR.
XX      The invention provides human NALAD-ase I, II and IV polypeptides,
XX      cDNAs, antisense nucleic acids, vectors, host cells, transgenic
XX      organisms, antagonists and agonists. These are useful for treating
XX      neural disorders such as Alzheimer's disease, schizophrenia, ALS,
XX      Parkinson's disease, peripheral neuropathy, Huntington's disease,
XX      acute brain injury, multiple sclerosis, exposure to neurotoxins,
XX      peripheral nerve trauma, ischemia or dementia (claimed). Nucleic
XX      acids can also be used for gene therapy and for genetic screening
XX      of predisposition to disorders associated with NALAD-ase.
SQ      Sequence 1884 BP; 520 A; 432 C; 458 G; 474 T; 0 other;
XX
XX
Alignment Scores:
Score: 2.15e-236 Length: 1884
Percent Similarity: 2417.00 Matches: 472
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
DB: Query Match: 100.00% Indels: 0
Gaps: 0
US-09-745-763-36 (1-472) x AAZ58313 (1-1884)
Qy      1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuSerLeuCysSerGly 20
Db      149 ATGAAATTCCTTATCTTGCATTTTTCGGTGGTTCACCTTATATCCCTGCTGGG 208
Qy      21 LysAlaIleCysLysAenGlyIleSerLysArgThrPheGlnGlnIleLysGlnGlnIle 40

```

```

Db      209 AAACCTATGTCAGAAATGCGATCTCTAAGAGGACCTTTGAAGAAATTAAGAGAAATA 268
Qy      41  AlaserCysglYaspValAlAlaYsAlaIleIleasneuAlaValTYrGlyValAlaGln 60
Db      269 GCCAGCTGTGGAGATGTTGCTTAAGCAATCATCAACCTAGCTGTTTATGTTAAGGCCAG 328
Qy      61  AsnaRserTYrGluArgLeuAlaleuLeuValAspThrValGlyProArgLeuSerGly 80
Db      329 AACAGATCTATGAGCGATTTGGCACTTCTGGTTGATATGTTGGACCCAGATGAGTGGC 388
Qy      81  SerLYsasnLeuGlyValAlaIleGlnIleMetTYrGlnAsnLeuGlnGlnAspGlyLeu 100
Db      389 TCCAAGAACCTAGAAAAGCCATCCAAATYATGTACAAAACCTGCACAGAAATGGGCTG 448
Qy      101 GlnLYsValHisLeuGlnProValArgIleProHisTPGluArgGlyGlnSerIle 120
Db      449 GAGAAAGTTCACCTGGAGCCAGTGAATACCCCACTGGAGAGGGAGAAATCAGCT 508
Qy      121 ValMetLeuGluProArgIleHisLYsIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db      509 GTGATGCTGGAGCCAGAAATTCATAGATACCATCTCGGCTTTGGCAGCAGCATTTGG 568
Qy      141 ThrProGluGlyIleThrAlaGlyValLeuValValThrSerPheAspGluLeuGln 160
Db      569 ACTCCTCCAGAGGATTACAGACAGAAAGTTCTGGGTGAGACTCTTTGATGAACTGCAG 628
Qy      161 ArgArgAlaSerGluAlaArgGlyLYsIleValValTYrAsnGlnProTYrIleAsnTYr 180
Db      629 AGAAGGGCTCTAGAGCAAGAGGAGGAAAGTTGTTGTTTAAACCAACCTTACATCAACATC 688
Qy      181 SerArgThrValGlnTYrArgThrGlnGlyValAlaGlyAlaAlaLYsValGlyValLeu 200
Db      689 TCAAGACGGTGCATATCCAGACGACGAGGGGGCGGTGGAGACTGCCAGAGTTGGGCTTTG 748
Qy      201 AlaSerLeuIleArgSerValAlaSerPheSerIleTYrSerProHisThrGlyIleGln 220
Db      749 GCATCTCTCATTCATCCGTGGCTCTTCTCCATCTACAGTCTCACACAGATATTCAG 808
Qy      221 GlnTYrGlnAspGlyValProLYsIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      809 GAAATCCAGATGGCGTCCAGAAATTCCAACGCTTATTAACGTTGAGAAATGCCAGAA 868
Qy      241 MetMetSerArgMetAlaSerHisGlyIleLYsIleValIleGlnLeuLYsMetGlyVal 260
Db      869 ATGATGTCAAAATGGCTTCTCATGGGATCAAAATGTCATTCACCTTAAAGTTGGGGCA 928
Qy      261 LYsThrTYrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLYsTYr 280
Db      929 AAGACCTAACCAAGATACGATTCCTTCAACACTGTAGCAGAGATCATCTGGGAGCAAAATAT 988
Qy      281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTPAspValGlyGlnGlyVal 300
Db      989 CCAAGAACGGTGTGAATCGTCACTGACATCTGCACACTGGAGTTGGCGAGGGTCC 1048
Qy      301 MetAspAspGlyGlyValAlaPheIleSerTPGluAlaLeuSerLeuIleLYsAspLeu 320
Db      1049 ATGGATGATGGGGGTGGAGCCCTTATATCATGGAGGACCTCTCATATTAAATATCTT 1108
Qy      321 GlnLYsAspProLYsArgTYrThrLeuArgLeuValLeuTYrThrAlaGluGlnGlnGly 340
Db      1109 GGGGCTGCTCCAAAGAGGACTCTGGGCTGGCTGTGCTGCATCGCAGAAAGAACAGTGGGA 1168
Qy      341 ValGlyAlaPheGlnTYrTYrGlnLeuHisLYsValAsnIleSerAsnTYrSerLeuVal 360
Db      1169 GTTGTGTGCTTCCAGTATTAATCACTTACCAAGGTAATATTTCCAACTACAGTCTGGTG 1228
Qy      361 MetGluSerAspAlaGlyThrPheLeuProThrArgLYsLeuGlnPheThrGlySerGlyLYs 380
Db      1229 ATGGAGTCTGACGCGAGAACTCTTACCACTGGGCTGCATTTACGCGCAGTAAAG 1288
Qy      381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400

```

```

Db      1289 GCCAGGCCCATCATGAGGAGGATTATGAGCTCTGTCAGGCCCTCAATATCATCAGTTC 1348
Qy      401  LeuSerHisGlyGlyGlyTYrThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyVal 420
Db      1349 CTGAGCCATGAGAGGAGGACAGACATCACTTTTGGATCCAAAGCTGGAGTGGCTGGAGCC 1408
Qy      421 SerLeuLeuAspAspLeuTYrLYsTYrPhePhePheHisHisSerHisGlyAspThrMet 440
Db      1409 AGTCTACTTATGATCATATACAGATATTTCTTTCATCATCACTCCACGGAGACACCATG 1468
Qy      441 ThrValMetAspProLYsGlnMetAsnValAlaAlaValTrpAlaValSerTYr 460
Db      1469 ACTGTCATGATCCAAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1528
Qy      461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db      1529 GTTGTGACGACATGAGAAATGCTCCTGAGTTC 1564

RESULT 9
AAZ98139
ID  AAZ98139 standard; cDNA; 1923 BP.
AC  AAZ98139;
DT  11-MAY-2000 (first entry)
DE  Human signal peptide containing protein HSP-31 cDNA SEQ ID NO:165.
XX  Human; signal peptide-containing protein; HSP; diagnosis; cancer;
KW  inflammation; cardiovascular disease; anticancer; anti-inflammation;
KW  antimicrobial; neuroprotective; cardioprotective; hepatocytic;
KW  antiasthmatic; gene therapy; cell proliferation; neurological disorder;
KW  reproductive disorder; developmental disorder; arteriosclerosis;
KW  cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
KW  asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
KW  Parkinson's disease; Huntington's diseases; ovulatory defect;
KW  muscular dystrophy; ss.
XX  Homo sapiens.
OS  Homo sapiens.
PN  WO200000610-A2.
PD  06-JAN-2000.
PE  25-JUN-1999; 99WO-UG14484.
XX  26-JUN-1998; 98US-0090762.
PR  31-JUL-1998; 98US-0094983.
PR  01-OCT-1998; 98US-0102686.
PR  11-DEC-1998; 98US-0112129.
XX  (INCY-) INCYTE PHARM INC.
PA  Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR,
PI  Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL,
PI  Bandman O;
XX  WP1; 2000-160673/14.
DR  P-PSDB; AAY87254.
XX  New human signal peptide-containing proteins useful in treatment,
PT  prevention and diagnosis of e.g. cancer, inflammation and
PT  cardiovascular disease
XX  Claim 9; Page 269-270; 327pp; English.
CC  AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the
CC  human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have
CC  anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatocytic,
CC  neuroprotective, cardioprotective, and antiasthmatic activities, and can
CC  be used in gene therapy. HSPs can be used to treat or prevent disorders
CC  associated with decreased activity or function of HSP. Antagonists of
CC  HSP are used to treat or prevent disorders associated with increased

```


CC activity or function of HSP. Such diseases include cell proliferation
CC (including cancer), inflammation, cardiovascular, neurological,
CC reproductive or developmental disorders, (e.g. arteriosclerosis,
CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
CC asthma, Crohn's disease, microbial or other infections, congestive or
CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP
CC nucleic acids can be used for the recombinant production of HSP, for
CC detecting HSP in standard hybridisation and amplification assays (for
CC diagnosis and monitoring), in gene therapy, as antisense,
CC triplex-forming or ribozyme therapeutics, for detecting related sequences
CC or genetic variations, and for chromosomal mapping. HSP are also used to
CC raise specific antibodies (Ab) and to screen for agonists and
CC antagonists (potential therapeutic agents). Ab are used to diagnose, or
CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic
CC antagonists, in competitive drug screens, and for purification of HSP
CC from natural sources.

XX Sequence 1923 BP, 538 A; 439 C; 471 G; 475 T; 0 other;

Alignment Scores:

Pred. No.:	2,21e-236	Length:	1923
Score:	2417.00	Matches:	472
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-745-763-36 (1-472) x AAZ98139 (1-1923)

Oy 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20
Db 179 ATGAATTCCTTATCTTCGATTTTCGGTGTTTCACTTTATCCCTGCTGCTGG 238
Oy 21 LysAlaIleCyLeuAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40
Db 239 AAAGCTATATGCMAAATGGCATCTCTTAAGAGACTTTTGAAGAAATAAAGAAATA 298
Oy 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
Db 289 GCCACTGTGAGATGTTGCTTAAGCATCATCACTGCTTTATGTGTAACCCAG 358
Oy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db 359 AACGATCTCTATGACGATGGCACTTCTGTTGATCTGTGGACCCAGACTGAGTGGC 418
Oy 81 SerLysAsnLeuGluLysAlaIleGlnIleLeuTyrGlnAsnLeuGlnAspGlyLeu 100
Db 419 TCCAAAGACTTAAGAAAGCCATCCAAATTATGTACCAAAACCTCAGCAAGATGGCTG 478
Oy 101 GluLysValHisLeuGluProValArgIleProHisTTPGluArgGlyGluGluSerAla 120
Db 479 GAGAAAGTTCACCTGAGCCAGTGAAGAAATCCCACTGGAGAGGGGAGAAAGAAATCAGCT 538
Oy 121 ValMetLeuGluProArgGlyIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 539 GTGATGCTGGAGCCAGAAATTCATTAAGATAGCCATCCTGGGCTTGGCAGCATTTGG 598
Oy 141 ThrProProGluGlyIleThrAlaGluValLeuValThrSerPheAspGluLeuGln 160
Db 599 ACTCTCTCAGAGAGGATTAACAGCAAGATTCGTGGTGCACTCTTTCGATCAACTGAG 658
Oy 161 ArgArgAlaSerGluAlaArgGlyLysIleValIleTyrAsnGlnProTyrIleAsnTyr 180
Db 659 AGAAGGGCTCAGAAAGCAAGAGGAAAGATGTTGTTTAAACCAACTTACATCACTAC 718
Oy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValAlaLysValGlyValLeu 200
Db 719 TCAAGCAGGCTGCAATACGAGACGAGGGGGCGTGGAAAGCTGCCAAGTTGGGCTTTG 778
Oy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db 779 GCATCTCATTCGATTCGCTGGCTCTCTTCATCTACAGCTCACACAGATATTGAG 838

Oy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 839 GAATACCAAGATGGCGGCCCAAGATTCCAAGACCTGTTATTCGGTGAAGATGCAAGAA 898
Oy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyVal 260
Db 899 ATGATGTCAAGAAATGGCTTCTCATATGGATCAAAATTCATTCATTCAGCTAAAGATGGGCA 958
Oy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
Db 959 AAGACTTACCCAGTACTGATTCCTTAAACACTGTACAGAGATCACTGGAGCAAAATAT 1018
Oy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyVal 300
Db 1019 CCAAGACAGGTGTACTGCTGCTGAGACATCTGACAGCTGGAGTGTGGCCAGGGTGGC 1078
Oy 301 MetAspAspGlyGlyValAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
Db 1079 ATGGATGATGGCGGTGAGCCTTATATCATGGAGACACTCTCATATTAAAGATCTT 1138
Oy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrPheAlaGluGluGlnGly 340
Db 1139 GGGCTGGCTCCAAAGAGAGACTGCGCTGCTGCTGAGCTGAGCAAGAAAGAGTGA 1198
Oy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
Db 1199 GTTGCTGCTTCCAGTATTCATGATTCACAGGTAATATTCATTCACATCAGCTGGTG 1258
Oy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
Db 1259 ATGAGATCTGACGAGAGAACCTTCTTACCACTGGCTGCAATTCACCTGGCAGTGAAG 1318
Oy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlyVal 400
Db 1319 GCCAGGCGCATGAGAGAGGTATAGCTGTGAGCCCTCAATATCATCACTAGGTC 1378
Oy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyVal 420
Db 1379 CTGAGCCATGAGAGAGGAGACAGCATCACTTTGGATCCAGCTGGAGTGGCTGGAGCC 1438
Oy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
Db 1439 AGCTACTTGATGACTTATCAAGATATTTCTTCCATCATCTCCACGGAGACACCATG 1498
Oy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValAlaThrPalAlaValSerTyr 460
Db 1499 ACTGTCAATGATCCAAAGCAGATGAATGTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1558
Oy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db 1559 GTTGTTCAGACATGAGAAAGAAATGCTGCTAGTCC 1594

RESULT 10
ID ACCS0603 standard, cDNA; 2077 BP.
XX ACCS0603;
XX 12-JUN-2003 (first entry)
DE Human secreted protein coding sequence, SEQ ID 270.
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyrostatic;
KW vulnerability; antiinflammatory; nootropic; neuroprotective;
KW antiparkinsonian; gene therapy; human; cardiovascular disorder;
KW gene; ss.
OS Homo sapiens.
XX Homo sapiens.
XX WO200295010-A2.
XX 28-NOV-2002.

XX 19-MAR-2002; 2002WO-US09785.
 PF 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM;
 DR WPI, 2003-129429/12.
 XX Novel human secreted proteins, useful for detecting, preventing,
 PT diagnosing, prognosticating, treating and/or ameliorating
 PT cardiovascular disorders such as arrhythmia -
 XX Claim 21, SEQ ID 270; 1681bp; English.
 PS
 CC The present invention relates to novel human secreted proteins
 CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The
 CC proteins and their coding sequences are useful for the preparation of a
 CC diagnostic or pharmaceutical composition for diagnosing or treating a
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
 CC coronary arteriosclerosis and myocardial ischemia), neural disorders,
 CC immune system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of
 CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism.
 CC Note: The sequence data for this patent was published in electronic
 CC format and is available from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,46e-236 Length: 2077
 Score: 2417.00 Matches: 472
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 25 Gaps: 0
 US-09-745-763-36 (1-472) x ACC50603 (1-2077)
 QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20
 DB 132 ATGAAATTCCTTATCTTGCAATTTTCGGTGCTGTCACCTTTATCCCTGTGCTGGG 191
 QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
 DB 192 AAACCTATATGCAAGATGCGATCTCTAGAGGACTTTTGAAGAAATTAAGAAAGAAATA 251
 QY 41 AlaSerCyGlyAspValAlaLysAlaIleIleLeuLeuAlaValIYrGlyLysAlaGln 60
 DB 252 GCCAGCTGTGAGATGTTGTCTTAAGACATCATCAACCTAGCTGTATGTGTAAGGCCAG 311
 QY 61 AsnArgSerTYrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
 DB 312 AACGATCTCTATGAGCGATGTGCACTTCTGTGTACTGTGTGACCCGACGACTAGTGGC 371
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTYrGlnAsnLeuGlnGlnAspGlyLeu 100

DB 372 TCCAAGAACCTAGAAAAAGCCATCAATTTGTACCAAAACCTGCAGCAAGATGGGCTG 431
 QY GluLysValHisLeuGluProValArgIleProHisTrpPheGluArgGlyGluSerAla 120
 DB 432 GAGAAAGTTCACCTGTGAGCCAGTAGAATATCCCATGGGAGAGGGGAGAAATCAAGCT 491
 QY ValMetLeuGluProArgIleHisIleLysIleAlaIleLeuGlyLysSerSerIleGly 140
 DB 492 GTGATGCTGAGCCCAAGATTCATTAAGATAGCCATCTCGGGTCTTGGCAGCAGCATGGG 551
 QY 141 ThrProGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160
 DB 552 ACTCCTCCAGAGAGCATATACGACGAGAGTTCTGTGTGAGACTCTTTCGATGAACCTGAG 611
 QY 161 ArgArgLAserGluAlaArgGlyLysIleValValIYrAsnGlnProTYrIleAsnTYr 180
 DB 612 AGAAGGGGCTCAGAGAGAGAGAGAGATTTGTTTATTAACCAACCTTACATCAACTAC 671
 QY 181 SerArgThrValGluTYrArgThrGlnGlyValAlaGluAlaLysValGlyValAlaLeu 200
 DB 672 TCAGAGACGTGCAATACCAACGAGGGGCGGTGAGAGCTGCCAAGTTGGGCTTTG 731
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTYrSerProHisTrpGlyIleGln 220
 DB 732 GCATCTCATTCATGCATCGGTGGCTCTCTTCATCACTACAGTCTTACACAGATTCAG 791
 QY 221 GluTYrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 DB 792 GAATACCAAGATGTGGCGCCCAAGATTCCAACACCTGTATTAACGTGGAAGATGCAGAA 851
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyVala 260
 DB 852 ATGATGTCAGAAATGGCTTCTCATGGAGTCAAAATGTGATTCAGCTTAAGATGGGGCA 911
 QY 261 LysThrTYrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTYr 280
 DB 912 AAGACTACCAAGATATCATGATCTCTTCAACCTGACAGATCACTGGAGCAAGATAT 971
 QY 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
 DB 972 CCAGAACAGGTGTATCGTGCAGTGCAGATCTGCAGACGTGGAGTGTGGCAGGGTGC 1031
 QY 301 MetAspAspGlyGlyValAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
 DB 1032 ATGATATATGGCGGTGAGGCTTTATATCATGGAGACACTCATCTTATTAAGATCTT 1091
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGluGlnGly 340
 DB 1092 GCGCTGCGTCCAAAGAGAGACTCTGCGGCTGTGCTGTGAGCTGAGAAAGAAAGTGGGA 1151
 QY 341 ValGlyAlaPheGlnTYrTYrGlnLeuHisLysValAsnIleSerAsnTYrSerLeuVal 360
 DB 1152 GTTGTGCTCTCCAGATTTATCAGTTACACAGAGTAATATTTTCCACTACAGCTGTG 1211
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
 DB 1212 ATGAGTCTGACGAGAGAACCTCTTACCACTGGGCTGCACATTCACGTGCAAGTGA 1271
 QY 381 AlaArgAlaIleMetGluGluValAlaMetSerLeuLeuGlnProLeuAsnIleThrGlyVal 400
 DB 1272 GCCAGGCGCATCATGAGAGAGGTTATAGCTGTGACGCCCTCATATATCATCAAGTGC 1331
 QY 401 LeuSerHisGlyGlyIYrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
 DB 1332 CTGAGCCATGAGAGAGAGAGACATCACTTTGTGATCCAAAGTGTGAGTGGCGAGGCC 1391
 QY 421 SerLeuLeuAspAspLeuTYrLysTYrPhePheHisHisSerHisGlyAspThrMet 440
 DB 1392 AGTCTACTGATGACTTATACAGATATTTCTTCTTCATCATCTCCACGAGACACATG 1451
 QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTrpAlaValIleSerTYr 460

Db 1452 ACTGTCATGATCCAAAGACATGATGCTGCTGTTGGCTGTTCTTAT 1511
Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db 1512 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1547
RESULT 11
AB271331
ID AB271331 standard; cDNA; 2077 BP.
AC AB271331;
XX
XX 04-APR-2003 (first entry)
DT
DE Secreted protein-encoding gene 142 cDNA clone HRA035, SEQ ID NO:152.
XX
XX Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
KW antiinflammatory; immunosuppressive; vulnereary; chromosome 8q22.2;
KW gene; ss.
XX
XX Homo sapiens.
OS
XX MO200276488-A1.
PN
XX 03-OCT-2002.
PD
XX 19-MAR-2002; 2002WO-US08276.
PF
XX 21-MAR-2001; 2001US-277340P.
PR 19-JUL-2001; 2001US-306171P.
XX 13-NOV-2001; 2001US-331287P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
XX Rosen CA, Ruben SM;
PI
XX WPI; 2003-029900/02.
DR P-PSDB; ABR00152.
XX
XX New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers -
XX
XX Claim 21; Page 841-842; 1216bp; English.
XX
XX AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
CC AB271479-AB271540 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening, and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing, treating, ameliorating or preventing
CC digestive disorders. Such conditions include disorders of the mouth,
CC oesophagus, stomach, small intestine, large intestine, liver, biliary
CC tract and pancreas, and include cancers of these organs and tissues. The
CC secreted proteins and their nucleic acids may also be used in the
CC treatment of immune disorders, inflammation, infection,
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
CC of the invention may be used for chromosome identification, chromosome
CC mapping, in gene therapy, for identifying individuals from minute
CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein-
CC encoding cDNA clone of the invention.
XX
XX Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;
SQ
Alignment Scores:

Pred. No.: 2,46e-236 Length: 2077
Score: 2417.00 Matches: 472
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Gaps: 0
Query Match: 100.00% Gaps: 0
US-09-745-763-36 (1-472) x AB271331 (1-2077)
Qy 1 MetLysPheLeuLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20
Db 132 ATGAAATTCCTTATCTTCGCGATTTTTCGGTGTTCACCTTTATCCCTGCTCGGG 191
Qy 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
Db 192 AAAGCTATATGCAAGAAATGGCATCTTAAAGAGCATTTTGAAGAAATTAAGAAATA 251
Qy 41 AlaSerCysGlyAspValAlaValAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
Db 252 GCCAGCTGTGGAGATGTTGGCTAAAGCAATCATCAACCTAGCTGTTATGTTAAAGCCGAG 311
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db 312 AACAGATCCTATATGAGCATTTGGCACTTCTGCTTATATCTGTGAGCCAGACTAGTGGC 371
Qy 81 SerLysAsnLeuGluLysValAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
Db 372 TCAGAAACCTTAAGAAAGCCATCCAAATATATGACAAACCTTGACAGCAAGATGGCTG 431
Qy 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluSerAla 120
Db 432 GAGAAAGTTACCTGGAGCCAGTGAGAAATCCCACTGGAGAGAGGAGAAATCAAGCT 491
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 492 GTGATGCTGAGCCAAATTCATAGATAGCCATCTCGGCTCTTGACAGCAGCATTTGG 551
Qy 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
Db 552 ACTCTTCAGAGGACATTTACGACAGAAATCTGTGGTGTACCTTTTCATGAACTGACG 611
Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
Db 612 AGAAGGCGCTCAGAAAGCAAGAGGAGATTTGTTATTAACAACCTTACATCACTAC 671
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaValValValGlyAlaLeu 200
Db 672 TCAAGGACGGTGCAATACCGAACGAGGGGGGTGAGAGCTGCCAAGTTGGGGCTTTG 731
Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db 732 GCATCTTCTCATTCATCCGCGGGCTCTTCTTCATCTACAGTCTTCACACAGATATTCAG 791
Qy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 792 GAATACAGAGATGGCGGCCAATTCGAACAGCCGTATTACGGTGAAGATGCGAGAA 851
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 852 ATGATGTCAGAAAGGCTTCTTCATGAGATCAAAATTCATTCAGCTAAAGATGGGGCA 911
Qy 261 LysThrTyrProAspThrArgSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
Db 912 AAGACCTTACCAAGATCTGATCTTCTTACACCTGTACAGAGATCATCTGGAGCAATAT 971
Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300
Db 972 CCAGAAACAGTTTACGTGCTAGTGACATCTGACACGCTGGATGTTGGCAGGGGCC 1031
Qy 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
Db 1032 ATGATGATGGCGGTGAGGCTTTATATCATGGAACACACTCTCATTAATAAGATCTT 1091

QY 321 GlyLeuArgProIleValGlnThrLeuArgLeuValLeuThrAlaGluGlnGly 340
DB 1092 GGGCTGGCTCCAAAGAGACTCTCGGCTGGCTGCTGACAGCAAGAGAGTGA 1151
QY 341 ValGlyAlaPheGlnIleTyrGlnLeuHisValAsnIleSerAsnTyrSerLeuVal 360
DB 1152 GTGGTCCCTCCAGTATTAACAGTTCACAAAGTAATATTTCCAACTACAGTCTGGTG 1211
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380
DB 1212 ATGAGAGCTAGACGAGAACCTTCTTACCCAGCTGGCTGCAATTCACCTGCGAGTGAAG 1271
QY 381 AlaArgAlaIleMetGluGlnValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
DB 1272 GCCAGGCCCATCAGAGAGAGTATGAGCTGCTGACGCCCTCAATATCATCTCAGGTC 1331
QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyVal 420
DB 1332 CTGAGCCATGAGAGAGGACAGACATCACTTTGGATCCAAAGCTGAGTCCCTGAGGCC 1391
QY 421 SerLeuLeuAspAspLeuTyrLeuTyrPhePhePheHisSerHisGlyAspThrMet 440
DB 1392 AGTCTACTGATGACTTATACAGATATTTCTTCCATCATCTCCACGAGACACCATG 1451
QY 441 ThrValMetAspProIleGlnMetAsnValAlaAlaValAlaValSerTyr 460
DB 1452 ACTGTCATGATCCAAAGACAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1511
QY 461 ValValAlaAspMetGluGlnMetLeuProArgSer 472
DB 1512 GTTGTTCAGACAGAGAAATGCTGCTAGGTC 1547

RESULT 12

AAA40493 standard; cDNA; 1767 BP.

AAA40493;

16-NOV-2000 (first entry)

Human fetal kidney cDNA fragment AM282_11.

Secreting protein; cytosolic; immunostimulatory; antimicrobial;
antiviral; immunosuppressive; antiinflammatory; vulnary; cytokine;
cell proliferation; differentiation; regulator; treatment; tumor;
autoimmune disease; inflammatory disorder; wound; microbial infection;
viral disease; graft versus host reaction suppression; ss.

Homo sapiens.

WO200037630-A1.

29-JUN-2000.

22-DEC-1999; 99WO-US31005.

23-DEC-1998; 98US-0220876.

(GBMY) GENETICS INST INC.

Jacobs K, McCoy JM, LaValle ER, Collins-Racie, LA, Evans C,

Merberg D, Treacy M, Bowman MR;

WPI; 2000-442661/38.

P-PSDB; AAB10229.

Secreted human proteins AS296-1i and AS34-1i, useful for treating

tumors, autoimmune diseases, inflammatory disorders, wounds, microbial

infections and viral diseases -
Disclosure; Page 198; 293pp; English.

This invention describes novel secreted human proteins (1) which have

CC cytosolic, immunostimulatory, antimicrobial, antiviral,
CC immunosuppressive, antiinflammatory and vulnary activity and which act
CC as cytokine, cell proliferation or differentiation regulators. (1)
CC is useful for treating tumors, autoimmune diseases, inflammatory
CC disorders, wounds, microbial infections and viral diseases. (1) is also
CC useful for suppressing graft versus host reaction. AAA40490-A40580
CC represent cDNA fragments that encode the secreted proteins
CC AAB10226-B10288 described in the method of the invention.

XX Sequence 1767 BP; 512 A; 384 C; 418 G; 453 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	1767
Score:	2410.00	Matches: 471
Percent Similarity:	99.79%	Conservative: 0
Best local Similarity:	99.79%	Mismatches: 1
Query Match:	99.71%	Indels: 0
DB:	21	Gaps: 0

US-09-745-763-36 (1-472) x AAA40493 (1-1767)

QY 1 MetLeuPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCysSerGly 20
DB 6 ATGAATTCCTTATCTTGGCATTTTCGAGGTTCACCTTTATCCCTGCTCTGGG 65
QY 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40
DB 66 AAGCTATATGCAAGATGACATCTCTAGAGACCTTTTGAAGAAATTAAGAAATA 125
QY 41 AlasSerCysGlyAspValAlaValAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
DB 126 GCCAGCTGTGAGATGTTGCTTAAGCAATCATCACTACTGTTTATGTTAAAGCCAG 185
QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
DB 186 AACAGATCTTATAGAGGATGCGATTCCTGTTGATATCTGTGACCCAGCATGAGTGC 245
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
DB 246 TCCAGAACCTTAAGAAAGCAATCCAAATATATGTAACAAACCTGACGAGATGGCTG 305
QY 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluIleSerAla 120
DB 306 GAGAAAGTTACCTGAGCCAGAGAGATATCCCACTGGAGAGGGAGAGATCGCT 365
QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
DB 366 GTGATGCTGAGGCAAGAAATTCATTAAGATAGCCATCTGGGCTTGGCAGCAGATTGGG 425
QY 141 ThrProGluGluIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160
DB 426 ACTCTCTCAAGAGCATTAACGAGAGATCTGAGTGTGACCTTTCAGAGAACTGAG 485
QY 161 ArgArgLaserGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
DB 486 AGAAGGCTCTCAGAGAGAGAGAGATGTTGTTTATAACCACTTATCATCACTAC 545
QY 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaAlaLysValGlyAlaLeu 200
DB 546 TCAAGGACGCTGCAATACCAACGAGGAGGCGGTGGAACCTCCAAAGGAGGAGCTTGG 605
QY 201 AlasSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisTyrIleGln 220
DB 606 GCATCTCTCATGATCGTGGCTCTCTTCCATCTACAGTCTCCACAGAGATTACG 665
QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
DB 666 GAATACAGAGATGGCGGCCCAAGATTCCAAAGCGCTGTTACGGTGGAGAGATGAGAA 725
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
DB 726 ATGATGCAAGAAATGCTCTCATGAGGATCAAAATGTCATTCAGCTAAAGATGGGGCA 785


```

Db 377 GTGATGCTGGAGCCAAAGATTTCATAGATGACATCCGCGCTTGGCAGCAGCATTTGGG 436
Qy 141 ThrProGluGlyIleThrAlaGluValIleThrSerPheAspGluLeuGln 160
Db 437 ACTCTCCAGAAAGCATTACACAGAGATTCTGGTGGAGCCTCTTGGATGAATGCGAG 496
Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValIleYaaGlnProTyrIleAsnTyr 180
Db 497 AGAAGGGCTCAGAAAGCAAGAGGGAAGATTGTTTAAACCAACTTACATCACTGAC 556
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200
Db 557 TCAAGGACGGTGCATATACGAAACGAGGGGGGGGGAAGCTGCCAAGGTGGGGCTTGG 616
Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db 617 GCATCTCTCATTCGATCCGTCGCTCTTCTCCATCTACAGTCTCCACACAGGATTTGAG 676
Qy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 677 GAATACCGAGATGGGCTGCCAAGATTCACACAGCTGATTTACGGTGGAAAGATCAGAA 736
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 737 ATGATGTCAAGAAATGGCTTTCATGAGATCAAAATTTGCATTCACACTAAAGATGGGGCA 796
Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280
Db 797 AAGACCTACCAAGATACGATTCCTTCAACACTGTAAGAGATCAGGGAGCAAAATAT 856
Qy 281 ProGluGlnValIleLeuValSerGlyHisLeuAspSerTyrPaspValGlyGlnGlyAla 300
Db 857 CCAAGAACGCTTGTACTGTCAGTGCATCTGCACTGCACTGGGATGTTGGGCGAGGTGCC 916
Qy 301 MetAspAspGlyGlyValAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
Db 917 ATGATGTATGGCGGGGAGCTTATATATCATGGGAAGCACTCTCATTTAAAGATCTT 976
Qy 321 GlyLeuAspProLysArgThrLeuArgLeuValLeuTyrThrAlaGluGlnGlyGly 340
Db 977 GGGGCTGCTCCAAAGAGACTCTGGGCTGGCTGCTGACAGCAGGAGAAACAAGTTGGA 1036
Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
Db 1037 GTTGGTGCCTTCCAGTATTTATCACTTACCAAGGAAATATTTCCAACTACAGTCTGGTG 1096
Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380
Db 1097 ATGAGCTGTAAGGAGCAACCTTCTTACCACTGGGCTGCAATTCACCTGGCGAGTAAAG 1156
Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db 1157 GCCAGGGCCATCATGAGAGAGGTTTGAAGCCCTGCGCAGCCCTCAATATCATCACTGAGTTC 1216
Qy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTyrIleGlnAlaGlyValProGlyAla 420
Db 1217 CTGAGCCATGAGAAAGGAGCAATCATCTTTGATTCAGAGCTGAGTGCCTGGAGCC 1276
Qy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisIleSerHisGlyAspThrMet 440
Db 1277 AGTCACTTGAATGATTTATACAGATATTTCTTCCATCACTCCACAGGAGACCACTG 1336
Qy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValTyrAlaValSerTyr 460
Db 1337 ACTGCATGAGATCCAAACAGATGATGTGCTGCTGTTGGCTGTTGTTGTTCTTAT 1396
Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db 1397 GTTGTTCAGACATGAAAGAAATGCTGCTAGGTCC 1432

```

RESULT 14
AAK94491

```

ID  AAK94491 standard; cDNA; 1860 BP.
XX
AC  AAK94491;
XX
DT  06-NOV-2001 (first entry)
XX
DE  Human full-length cDNA, SEQ ID NO: 3328.
XX
KW  Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS  Homo sapiens.
XX
PN  EPI130094-A2.
XX
PD  05-SEP-2001.
XX
PF  07-JUL-2000; 2000EP-0114089.
XX
PR  08-JUL-1999; 99JP-0194486.
XX
PR  11-JAN-2000; 2000JP-0118774.
XX
PR  02-MAY-2000; 2000JP-0183765.
XX
PA  (HELI-) HELIX RES INST.
XX
PI  Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX  Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX  WPI; 2001-524255/58.
XX  P-PSDB; AAM93559.
XX
PT  830 Primers useful for synthesizing full length cDNA clones and their
PR  use in genetic manipulation -
XX
ES  Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.
XX
CC  The invention relates to primers for synthesizing full length cDNA
CC  clones. 830 cDNA molecules encoding a human protein have been
CC  isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC  molecules have been determined. Primers for synthesizing the full length
CC  cDNA are useful for clarifying the function of the protein encoded by
CC  the cDNA. The full length clones were obtained by construction of full
CC  length enriched cDNA libraries that were synthesized by the oligo-capping
CC  method. The primers enable the production of the full length cDNA easily
CC  without any special methods. The present sequence is a full length
CC  human cDNA of the invention.
CC  Note: The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ  Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;
XX

```

Alignment Scores:

Pred. No.:	Length:	1860
Score:	2402.00	Matches: 469
Percent Similarity:	99.79%	Conservative: 2
Best Local Similarity:	99.36%	Mismatches: 1
Query Match:	99.38%	Indels: 0
DB:	22	Gaps: 0

```

US-09-745-763-36 (1-472) x AAK94491 (1-1860)
Qy 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuSerLeuCysSerGly 20
Db 129 ATGAAATTCCTTATCTTCGCAATTTTGGTGGTGTTCACCTTTATCCCTGCTGGG 188
Qy 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
Db 189 AAGCTATATCAAGATGCGATCTTAAGAGGACTTTTGAAGAATAGAGAAATA 248
Qy 41 AlaSerCysGlyAspValAlaLysAlaIleLeuAsnLeuAlaValTyrGlyLysAlaGln 60
Db 249 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCAACCTAGCTGTTTATGAGTAAAGCCAG 308
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLysSerGly 80

```



```

Db      ||| 309 AACAGATCTCATGACGCGATTGGCAGCTTCGGTTGATACGTGTGACCCAGACGAGTGAAGTGGC 368
Oy      81 SerIysAsnLeuGluValAlaIleGlnIleMetTyrGlnAsnLeuGlnIAspGluValLeu 100
Db      369 TCCAAAGACCTTGAAGAAAGCCATCCAAATTAATGATGACAAACCTTCAGCAAGATGGGCTG 428
Oy      101 GluIysValHisLeuGluProValArgIleProHisSTTPGluArgGluValGluIleSerAla 120
Db      429 GAGAAAGTTTCACTGAGCGACGACGATGAGAAATCCCACTGGAGAGGGGAGAGAAATCACT 488
Oy      121 ValMetLeuGluProArgGluIleHisIleValAlaIleLeuGluValSerIleGly 140
Db      489 GTGATGTGGAGCCAAAGAAATTCATAAGATAGCCATCTGGGCTTGGCAGCAGCAGTGGG 548
Oy      141 ThrProProGluGluValIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160
Db      549 ACTCTCCAGAGAGGATTAACAGCAAGATTCGTGGTGAGACCTCTTTCGATGAACTGACG 608
Oy      161 ArgATGAlaSerGluAlaArgGluValIleValIleValAlaTyrAsnGlnProTyrIleAsnTyr 180
Db      609 AGAAGGGCTTCAGAAAGCAAGAGGAGAGATTTGTTTATTAACCACTTACATCACTAC 668
Oy      181 SerArgThrValGlnTyrArgThrGlnGluValAlaGluAlaAlaIleValAlaValLeu 200
Db      669 TCAAGCAGGTGCATATACGAAACGAGGGGGCGGTGGAAGCTGCCAAGGTTGGGCTTGG 728
Oy      201 AlaSerIleIleArgSerValAlaSerPheSerIleTyrSerProHisSTHClGlyIleGln 220
Db      729 GCATCTTCATTCGATCCGCGTGGCTCTCTTCATCTCAAGCTTCACACAGAGATTCAG 788
Oy      221 GluTyrGlnAspGluValProIleValProThrAlaCysIleThrValGluIleAspAlaGlu 240
Db      789 GAATTCAGAGATGGGCTGCCCAAGATTCCAACGCTGATTAACGTTGAGAAATCCAGAA 848
Oy      241 MetMetSerArgMetAlaSerHisGlyIleIleValIleValIleGlnLeuIleMetGlyAla 260
Db      849 ATGATGTCAGAAATGCTTCATGAGGATCAAAATTTGATTCACATCAAGATGAGGGGCA 908
Oy      261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280
Db      909 AAGACCTTACCCAGATACGATTCCTTCAACACGTAAGAGATATCAGGAGGAGAAATAT 968
Oy      281 ProGluGlnValValIleuValSerGlyHisLeuAspSerTTPAspValGluIleValAla 300
Db      969 CCAAGAACAGTTTGTACTGTCAGTGCATCTGACAGCTGGAGATCTTGGGAGAGGTGCC 1028
Oy      301 MetAspAspGluGlyValAlaPheIleSerTTPGluAlaLeuSerLeuIleIysAspLeu 320
Db      1029 ATGATGATGGCGGTGGAGCCTTTATATCATGGAAGCACTCTCACTTATTAAGATCTT 1088
Oy      321 GlyLeuAspProIleValArgThrIleuArgLeuValIleuTTPThrAlaGluGluIleGly 340
Db      1089 GGGCTGCTCCAAAGAGACTCTGGCGGTGCTCTGAGACGCAAGAAACAAAGTGGGA 1148
Oy      341 ValGluAlaPheGlnTyrTyrGlnIleuHisIleValAlaAsnIleSerAsnTyrSerLeuVal 360
Db      1149 GTTGATGCTTCACATATTAATGATTAACAGAGTAATATTTCCACTACAGCTTGGTG 1208
Oy      361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluIys 380
Db      1209 ATGAGATCGAGCAGCAAGACTTTTATCCCACTGGGCTGCATTCACAGCGAGTGAAGA 1268
Oy      381 AlaAlaGluAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1269 GCCAGAGGCTATCGTAGAGAGGTTTAGCCCTGCGACGCCCTCAATATCACTACAGGTTC 1328
Oy      401 LeuSerHisGlyGluGlyThrAspIleAsnPheTTPIleGlnAlaGlyValProGluValAla 420
Db      1329 CTGAGCCATGAGAGAGGACAGACATCACTTTTGGATTCAGACCTGAGAGTGGCTTGAAGCC 1388
Oy      421 SerLeuLeuAspAspLeuTyrIleTyrPhePhePheHisHisSerHisGlyAspThrMet 440

```

```

Db      1389 AGTCTACTGATGACTTATACAGTATTTCTTTCATCATCACTCCACGAGAGACCATG 1448
Oy      441 ThrValMetAspProIleGlnMetAsnValAlaAlaIleValTTPAlaValIleSerTyr 460
Db      1449 ACTGTCATGATCCAAAGACATGATATGTTCTCTGCTGTGTTGGCTGTTGTTCTTAT 1508
Oy      461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db      1509 GTTGTCAGACATGAGAAAGAAATGCTGCTAGTGC 1544

RESULT 15
AAH9703 standard; cDNA; 1895 BP.
AAH9703;
16-OCT-2001 (first entry)
Human protein encoding cDNA sequence SEQ ID NO:538.
Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
antibacterial; antirheumatic; antiarthritic; immunosuppressive;
anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anaemia;
antiallergic; haemostatic; vulnery; antileuk; osteopathic; eczema;
dermatological; antiallergic; antidiabetic; antidiabetic; cytostatic;
neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
genetic disease; haematopoietic disorder; platelet disorder; asthma;
thrombocytopenia; osteoporosis; severe combined immunodeficiency;
allergic rhinitis; diabetes; multiple sclerosis; depression;
Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
neurological disorder; ss.
Homo sapiens.
MO200153455-A2.
26-JUL-2001.
22-DEC-2000; 2000WO-US35017.
23-DEC-1999; 99US-0471275.
21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Dymnac RT;
WPI; 2001-457603/49.
P-PSDB; AAM25762.
Isolated human polynucleotides encoding polypeptides, useful for the
treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
Claim 1; Page 591; 1217pp; English.
AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
AAM25963. The proteins can have activities based on the tissues and
cells they are expressed in, such as: antineoplastic; antirheumatic;
antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
central nervous system; virucide; anti-HIV; fungicide; antimutagen;
cardiovascular; antianemic; antiallergic; haemostatic; vulnery;
antileuk; osteopathic; dermatological; antiallergic; antidiabetic;
antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
antiparkinsonian; and immunostimulant. The proteins and polynucleotides
encoding them can be used in gene therapy, antisense therapy and vaccine
production. The proteins and polynucleotides are useful for screening for
agonists or antagonists of a protein and for the treatment and diagnosis
of disorders associated with the activity of a protein e.g. inflammation,

```

CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, hematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.

XX Sequence 1895 BP: 530 A; 439 C; 450 G; 476 T; 0 other:

Alignment Scores:

Pred. No.:	1,74e-228	Length:	1895
Score:	2339.50	Matches:	468
Percent Similarity:	98.32%	Conservative:	0
Best Local Similarity:	98.32%	Mismatches:	4
Query Match:	96.79%	Indels:	4
DB:	22	Gaps:	1

US-09-745-763-36 (1-472) x AAH9703 (1-1895)

```

Qy 1 MetlyspheleuilepheAlaPhePheGlyGlyValHisleuLeuSerleuCySerGly 20
Db 126 ATGAATTCCTTATCTTGCGCATTTTGGTGTGTCACCTTTATCCTGTGCTGGG 185
Qy 21 LysAlaIleCylysAsnGlyIleSerlysrGThrPheGluGluIleGlyGluIle 40
Db 186 AAGACTATATCCAGAAATGCGATCTCTAAGAGCATTTTGAAGAAATTAAGAAATA 245
Qy 41 AlaserCyserGlyaspValAlaLysAlaIleleuLeuAlaValIYrglyLysAlaGln 60
Db 246 GCCAGCTGTGAGATGTTGCTTAAGCAATCATCACTAGCTGTTATGTGTAAGCCGAG 305
Qy 61 AsparGserTyrgluArgleuAlaLeuLeuValAspThrValIgyProArgLeuSerGly 80
Db 306 AACGATCTTAATGAGCCATTTGCACTTCTGTGATACCTGTGAGCCAGCTAGTGGC 365
Qy 81 SerlyAsnLeuGluLysAlaIleGlnIleMetTyrglnAsnLeuGlnAspGlyLeu 100
Db 366 TCCAGAACTCTAAGAAAGCCATCCAAATTTATGCCAAACCTGCAAGAGATGGGCTG 425
Qy 101 GluLysValHisleuGluProValArgIleProHisTrpGluArgIyGluGluSerAla 120
Db 426 GAGAAAGTCACTGAGCCAGTGAATATACCCCACTGGAGAGGGGAGAAATTCAGCT 485
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleleuGlyLeuGlySerSerIleGly 140
Db 486 GTGATGCTGAGCCAGAAATTCATAGATACCAATCCCTGCTTGGCAGCAGCATTGGG 545
Qy 141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
Db 546 ACTCTCTCAAGAGCATTTACAGAAAGTCTGTGGTGAACCTCTTTCGAGAACTGCAG 605
Qy 161 ArgArgAlaSerGluAlaArgIyLysIleValValIYraGlnProTyrlAsnTyx 180
Db 606 AGAAGGGCTCAAGAGCAAGAGGAGATTGTTTATTAACCAACTTATCATCACTAC 665
Qy 181 SerArgThrValGlnTyraGTh-GlnGlyAlaValGluAlaIalysValGlyAlaLeu 200
Db 666 TCAGAGGACGGTGAATACCGAACCAGGGGGGGGAGAACTGCCAAGTTGGGGCTTTG 725
Qy 201 AlaSerleuIleArgSerValAlaSerPheSerIleTyrserProHisThrGlyIleGln 220
Db 726 GCATCTCTCATTCATCCGTGGCTCTCTTCATCTACATCACTCTCAACAGATTTAG 785
Qy 221 GluTyrglnAspGlyValProLysIleProThrAlaCyAlleThrValGluAspAlaGlu 240
Db 786 GAATATCCAGATGGGCTGCCAAGATTCCAACAGCTGTATTAAGTGAAGATGCAGAA 845
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 846 ATGATGTCAGAAATGGCTTTCATAGGATCAAAATTTGTCATTCAGCTAAGATGGGGCA 905

```

```

Qy 261 LysThrTyProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerTyx 280
Db 906 AAGACCTACCCAGATTAAGTATTCCTTCAACACTGTAGCAGATCACTGGAGCAATAT 965
Qy 281 ProGluGlnValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
Db 966 CCAGAACAGGTTTACAGTGCAGTGCAGATCTGCAGAGCTGGAGATGTTGGGAGGGTGC 1025
Qy 301 MetAspAspGlyGlyGlyAlaPheIleSerTrpGluAlaLeuSerleuIleLysAspLeu 320
Db 1026 ATGAGTATGCGCGTGGAGCTTTATATCATGGAAGCACTTCATTTATTAAGATCTT 1085
Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGlyGly 340
Db 1086 GGGGTGGTCCAAAGAGCATCTGCGGCTGTCTCTGAGCTGCAGAGAAACAAGTGA 1145
Qy 341 ValGlyAlaPheGlnTyrglnLeuHisLysValAsnIleSerAntyrserLeuVal 360
Db 1146 GTTGTCCTTCCAGATTTATCATGTTACACAGGTAAATATTTCCAACTACAGTGTG 1205
Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
Db 1206 ATGAGTCTACCGAGAACCTTTATCCCACTGGGCTGCATTCAGTGCAGTGAAG 1265
Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuGlnProLeuAsnIleThrGlnVal 400
Db 1266 GCCAGGSCCATCATGAGAGGATTAAGCTCTGAGGCCCTCAATATCATCACTAGTC 1325
Qy 401 LeuSerHisGlyGlyGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
Db 1326 CTGAGCCATGAGAAAGGAGCAGACATCACTTTGGATCCAGCTGAGAGCTGGAGCC 1385
Qy 421 SerLeuLeuAspAspLeuTyrlvsTyrrPhePhePheHisSerHisGlyAspThrMet 440
Db 1386 AGTCTACTATGACCTTAACAAGTATTTCTTCCATCATCACTCCACGAGACACCATG 1445
Qy 441 ThrValMetAsp---ProLysGlnMetAsnVal-AlaAlaIalVal-TrpAlaValAls 459
Db 1446 ACTGTCAATGGATCCAAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTGTTT 1505
Qy 459 exTyrrVal-ValAlaAspMetGluGluMetLeuProArgSer 472
Db 1506 CTATATGTTGTGCAGACATGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1547

```

Search completed: December 22, 2003, 23:51:21
 Job time : 398 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 22, 2003, 23:13:56 / Search time 2345 Seconds
(without alignments)
4891.988 Million cell updates/sec

Title: US-09-745-763-36
Perfect score: 2417
Sequence: 1 MKELIFAFGFGVHLISICSG.....AWAVSVYVADMEMLPRS 472

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 segs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlp
-O=/cg2_1/USFPO.spool/p/US9745763/runat.22122003_113317_25262/app_query.fasta.1.647
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCTL=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=400 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09745763.@CCN.1.1.2874.@runat.22122003_113317_25262 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hci:*
9: gb_hci:*
10: gb_est2:*
11: gb_hci:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_phg:*
26: em_gss_vr1:*
27: em_gss_vr1:*
28: gb_gss1:*

29: gb_gss2:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2133	88.2	1739	11 AK032972	AK032972 Mus muscu
2	2133	88.2	1805	11 AK075686	AK075686 Mus muscu
3	2124	87.9	1868	11 BC010977	BC010977 Mus muscu
4	1633	67.6	1024	11 BC012019	BC012019 Homo sapi
5	1539	63.7	1050	13 BX439467	BX439467 BX439467
6	1431	59.2	1038	13 BX355940	BX355940 BX355940
7	1352	55.9	909	13 BQ878966	BQ878966 BX360507
8	1312.5	54.3	1126	13 BX360507	BX360507 BX360507
9	1301.5	53.8	902	10 BG751497	BG751497 602730076
10	1290	53.4	919	13 BX335995	BX335995 BX335995
11	1277	52.8	866	13 BU146905	BU146905 AGENCOURT
12	1253	51.8	739	14 CB958693	CB958693 AGENCOURT
13	1252	51.8	1012	13 BX416896	BX416896 BX416896
14	1244	51.5	996	13 BX355939	BX355939 BX355939
15	1213	50.2	926	13 BQ938234	BQ938234 AGENCOURT
16	1204	49.8	998	13 BX460463	BX460463 BX460463
17	1200	49.6	804	14 CD352906	CD352906 UI-M-GLO-
18	1190	49.2	3151	11 BC017373	BC017373 Mus muscu
19	1178.5	48.8	813	14 CD519131	CD519131 AGENCOURT
20	1166	48.2	708	14 CD366462	CD366462 UI-H-FT1-
21	1164.5	48.2	922	12 BG761741	BG761741 602717936
22	1154	47.7	945	12 BI909780	BI909780 603070868
23	1144.5	47.4	889	10 BE906771	BE906771 60148329
24	1130.5	46.8	1081	13 BQ072892	BQ072892 AGENCOURT
25	1112	46.0	1020	13 BX439466	BX439466 BX439466
26	1105	45.7	801	12 BI754143	BI754143 603025609
27	1092	45.2	701	12 BI771109	BI771109 60305986
28	1092	45.2	783	12 BI754468	BI754468 603022607
29	1086	44.9	650	10 BG432644	BG432644 60246315
30	1068	44.2	786	10 BG430966	BG430966 602500267
31	1063.5	44.0	944	12 BI819273	BI819273 603037713
32	1059	43.8	879	13 BQ218838	BQ218838 AGENCOURT
33	1053	43.6	653	10 BE218907	BE218907 hv46b06.x
34	1052.5	43.5	1073	13 BX360506	BX360506 BX360506
35	1043.5	43.2	822	12 BI911772	BI911772 603065377
36	1036.5	42.9	805	10 BG563740	BG563740 602584524
37	1031	42.0	677	10 BE670671	BE670671 7638a06.x
38	1014	42.0	635	14 CB448111	CB448111 702171 MA
39	1000	41.4	736	12 BI769575	BI769575 603054947
40	996	41.2	676	9 AA796802	AA796802 vp31d01.r
41	995	41.2	751	14 CD558931	CD558931 AGENCOURT
42	990	41.0	637	13 BQ564473	BQ564473 g117610.y
43	983.5	40.7	849	10 BG166321	BG166321 602345529
44	980	40.5	806	12 BI764052	BI764052 603043258
45	979	40.5	727	12 BG765433	BG765433 602738911

ALIGNMENTS

RESULT 1
LOCUS AK032972
DEFINITION Mus musculus 12 days embryo male wolffian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:672048C06
product:Plasma glutamate carboxypeptidase, full insert sequence.
ACCESSION AK032972
VERSION AK032972.1 GI:26328732
KEYWORDS HTCC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
99279253
10349636

1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
MEDLINE
PUBMED
20499374
1:042159

2
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Katsuna, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
20530913
1:076861

3
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Baralov, S., Casavant, T.,
Pletschmann, W., Gaasterland, T., Gissi, C., King, B., Kochwa, H.,
Knehl, P., Lewis, S., Matsuo, Y., Nakado, I., Pescio, G.,
Queckenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarini, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Pletcher, C., Fujita, M., Gariboldi, M., Gastreich, S., Hill, D.,
Hochmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Machinomi, L., Washima, J., Mazzarelli, J., Momberts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmberg, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, T., Kawaji, H., Kotsuki, S.,
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

JOURNAL
MEDLINE
PUBMED
21085660
1:1217851

4
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
MEDLINE
PUBMED
21085660
1:1217851

5
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hirozane, T.,
Hori, F., Imotani, K., Iehi, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome

COMMENT
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES
SOURCE
Location/Qualifiers
1..1739
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:6720483C06"
/db_xref="taxon:10090"
/clone="6720483C06"
/sex="male"
/tissue_type="olfactory duct includes surrounding region"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
81..1493
/note="unnamed protein product: plasma glutamate
carboxypeptidase (MCD) [MGI:1889205, GI|NM_018755, evidence:
BLASTN, 99%, match=943]
putative"
/protein_id="BAC28105.1"
/db_xref="GI:26328733"
/transcript="MUSLPLFIVHLALSGKAVPNCVSQRTFREKEIANEYV
AKA1INAVGKTONSYERLGLVDVGRSGSKULEPAIDIMONLQDGLBNH
LEQVRIPEHMGESAVMLPRIRKVAIILGSSIGTPPGITAEVLVVASFELOR
ASEARGLIYVNOPTGYEVQYQAVAEAKVGVASLIQSVSFSYSPHTGIC
KYDGVKGIPTACTIVDAEMSRGKVIYHLEMAKTYPDTSFNTVABITS
MYPEVAVISGHSWGDGALDDGAGATISWALSTVADIGLRPRTLRVLMTE
EKGICASQYELHKNKISKYSLYMEADSTPTPTGICPGSDKARIMKXVNLLOP
LNVATKVSNEGDIVNIGAGVGSALRDVLKYPFHSHSDTMTMPKQNVAA
AAWATVAAVNEADDEMIPRS"
1722..1727
/note="putative"
1739
/note="putative"
BASE COUNT 482 a 369 c 423 g 465 t
ORIGIN
polya_signal
polya_site
Alignment Scores:
Pred. No.: 1,47e-235 Length: 1739
Score: 2135.00 Matches: 414
Percent Similarity: 93.64% Conservative: 28
Best Local Similarity: 87.71% Mismatches: 28
Query Match: 88.25% Indels: 2
DB: 11 Gaps: 1
US-09-745-763-36 (1-472) x AK032972 (1-1739)
Qy 1 MetLysPheLeuIlePheAlaPheGlyGlyValHisLeuLeuSerLeuCySergly 20
Db 81 ATGAGGCCCTTTCTTCCTCGTTTCAT-----GTTACCTTTTAAAGCTGGCGCTCGT 134
Qy 21 LysAlaIleCySlySaanglyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
Db 135 AAACCTGTAATTCAGAAATGGTGTTCCACGCGAATTCGAGAAATTAAGAAAGAAATA 194
Qy 41 AlSerCySglYpValAlaIleValAlaIleLeuLeuValAlaIleValAlaGln 60
Db 195 GCCACATGATGAGGTGTGCTAAAGCAATATCACTTCCTGTTTATGTAATATACAG 254
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80

Db 255 AACGGTCTATGAGCTTTGGACCTTCACTGATATCTGTGGACCCAGACTGATGGC 314
 Qy 81 SerIysenleugluysalalegllelmetYrGlnsenleuglnglnaspGlyLeu 100
 Db 315 TCTAAGAACTAGAGAAAGCTATTCATATCATGTATACCAAAACCTGCAGCAAGATGGCTA 374
 Qy 101 GlnlyValHisleugluProvalArgIleProHisTrpGluArgGlyGlnSerAla 120
 Db 375 GAAATGTCACTGAGCAGGTGAGATATCCCACTGGAGAGGAGAGAAATCTGCA 434
 Qy 121 ValMetleugluProalglleHisIleValleuGlyLeuGlySerSerIleGly 140
 Db 435 GTGATGCTTACCTCCAAATTCACAAATGCTTATTTAGTCTTGGACACAGCATTTGG 494
 Qy 141 ThrProGluGlyIleThrAlaGluValleuValValThrSerPheaspGluLeuGln 160
 Db 495 ACTCCCAAGAGGACATCACAGCAAGAGTCTGTGTGAGCTTTTGTATGAATCTTCAA 554
 Qy 161 ArgArgAlaSerGluAlaArgGlyIleValValIleValIleValIleValIleVal 180
 Db 555 AGAAGACATCAGAAAGCAAGAGGAGATCATTTATATACAGCCTTACACTGGCTAT 614
 Qy 181 SerArgThrValGlnIleArgThrGlnGlyAlaValGluAlaAlaValGlyAlaLeu 200
 Db 615 GAGAAAGCTGTGCACTACCGGCTCAGGAGCTGTGGAAGCTGCAGAGTGGAGCTGTG 674
 Qy 201 AlaSerleuIleArgSerValAlaSerPheSerIleTyrSerProHisTrpGlyIleGln 220
 Db 675 GCATCCCTCATCATCATGATGCTTCTTTCATCATCACTCATCACTCATCACTCATCA 734
 Qy 221 GlnIleGlnaspGlyValProIleIleProThrAlaCysIleThrValGluaspAlaGlu 240
 Db 735 AAATATCAAGATGTGTGTGCCCAATTCACAAACGCTGTATCACAGTAAAGATGCA 794
 Qy 241 MetMetSerArgMetAlaSerHisGlyIleIleValIleValIleGlnLeuIleMetGlyAla 260
 Db 795 ATGATGCTCGAAAGCTTCTCTCGTGGAAACAAATTCATTCATTCGGAATGGAGCA 854
 Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280
 Db 855 AAGACCTATCCAGATCATGATCTTCTTCAATTCGTGTGCAAGATCATCACTGGAGCATGTAT 914
 Qy 281 ProGluGlnValIleValIleValIleValIleValIleValIleValIleValIleVal 300
 Db 915 CCGAAGAGAGTGTCTGTGTGAGGACATTTGCAAGCTGGATGTTGGCAGAGTGCA 974
 Qy 301 MetAspAspGlyGlyIleValIleValIleSerTrpGluAlaLeuSerIleIleYaspIleu 320
 Db 975 CTGGATGATGTGTGTGAGCTTCTCATATCATGGAAGCACTCATCTGTTAAAGATCTT 1034
 Qy 321 GlnIleuArgProIleArgThrIleuArgIleuValIleuTrpThrAlaGluGlnGlnGly 340
 Db 1035 GGGGTGCTGCCAAAGAGACTCTGCTGTGTCTGTGACTGCAAGAAACAAAGAGGA 1094
 Qy 341 ValIleAlaPheGlnIleTyrGlnIleuHisIleValAlaIleSerAsnTyrSerIleVal 360
 Db 1095 ATTGTGCTCTCCCAATATTAAGACTACATAAGCAATATTTCCAACTACACCTGTGTG 1154
 Qy 361 MetIleSerAspAlaGlyThrPheIleuProThrGlyLeuGlnIlePheThrGlySerGlu 380
 Db 1155 ATGGAAGCTGATCAGAAAGAGTATGATCTTCTGCAACCCCTCAAGTACCAAGATGC 1214
 Qy 381 AlaArgAlaIleMetGluGlnValIleMetSerIleuGlnIleProIleuAlaIleThrGlnVal 400
 Db 1215 GCCAGGGCTATCATGAAGAGATGATGATCTTCTGCAACCCCTCAAGTACCAAGATGC 1274
 Qy 401 LeuSerHisGlyGlnGlyThrAspIleAsnThrTrpIleGlnAlaGlyValProGlyAla 420
 Db 1275 TTTTGTATATGAGAAAGAACTGATTTAATCTTGCATCCAACTGAGATGCTGTGAGCC 1334
 Qy 421 SerIleuLeuAspAspIleuTyrIleTyrPhePhePheHisIleSerHisGlyIleAspThrMet 440
 Db 1335 AGTCTGCTGATGATCTTGTACAAAGTATTTCTTTTCCATCATCTCCCAAGAGACACCATG 1394

Qy 441 ThrValMetAspProIySGlnMetAsnValAlaAlaAlaValTrpAlaValIleSerTyr 460
 Db 1395 ACTGTATGATGATCCAAAGAGATGATGTTGCTGCTGCTATGAGGCTGTGTGCTTAT 1454
 Qy 461 ValValAlaAspMetGluGluMetIleuProAspSer 472
 Db 1455 GTTGTTCAGACATGATGATGAATGCTGCCAGGTCC 1490
 RESULT 2
 AK075686
 LOCUS
 DEFINITION
 Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:119003P12 product:plasma glutamate carboxypeptidase, full insert sequence.
 AK075686
 ACCESSION
 AK075686.1 GI:26344480
 VERSION
 KEYWORDS
 HTc; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636
 REFERENCE
 AUTHORS
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 20499374
 PUBMED
 11042159
 REFERENCE
 AUTHORS
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Niehi, K., Kitesuna, T., Tashtiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, K., Ohata, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, D., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 TITLE
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL
 Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE
 20530913
 PUBMED
 11076861
 REFERENCE
 AUTHORS
 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aikawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamahata, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsumoto, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Adono, H., Baldarelli, R., Baxeh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Guerinich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Nazarelli, J., Mommaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.
 TITLE
 Functional annotation of a full-length mouse cDNA collection
 JOURNAL
 Nature 409 (6821), 685-690 (2001)
 MEDLINE
 21085660
 PUBMED
 11217851

REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
TITLE Group Phase I & II Team
JOURNAL Analysis of the mouse transcriptome based on functional annotation
REFERENCE Nature 420, 563-573 (2002)
AUTHORS 6 (bases 1 to 1805)
Adachi, J., Atzawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1..1805
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:1190003P12"
/db_xref="MGI:1909852"
/db_xref="taxon:10090"
/clone="1190003P12"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18-day embryo"
147..1559
/note="unlabeled protein product; plasma glutamate
carboxypeptidase (MGP) MGI:1889205, GI:NM_018755, evidence:
BLASTN, 99%, match=943)
putative"
/codon_start=1
/protein_id="BAC35891.1"
/db_xref="GI:26344481"
/db_xref="MGI:1923585"
/translation="MRSLFLFLFYHLALSGKAVFKVGSQRTFRREKEIANEDV
LEQVAILPMERGESAVMLEPRHMLALIGSSIGTPGGITAEVLVASFDEIQR
ASEARGLIVNQPYTGKTYQYQVQAAKVAASLQVSAFSLSPSTGIG
KYODGPKIPRACTIVEDAEMMSRMAKGIIVHLEAGATVPTDSENFVAETGS
MYPERVTVSGHIDSDNYGCGALDGGCAFLSMELSLVLDGAPKRTLELVMTAE
EGGIGASQYIEHKAHNSKISLVMEADSGFLPTGLDPTSDKRAIKKVMNLDP
LNTKVFPSNGEGTDINFWIQGVPGASLRDLTKYFFPHSHGDTMTWPKQNVAA
AVMAVAVYVADMDMLPRSS"
polyA_signal
1788..1793
/note="putative"
1805
/note="putative"
BASE COUNT 486 a 397 c 445 g 476 t 1 others
ORIGIN
Alignment Scores:

Pred. No.: 1,55e-235 Length: 1805
Score: 2133.00 Matches: 414
Percent Similarity: 93.64% Conservative: 28
Best Local Similarity: 87.71% Mismatches: 28
Query Match: 88.25% Indels: 2
Gaps: 1
US-09-745-763-36 (1-472) x AK075666 (1-1805)
QY 1 MetlyspheueuilepheuallphepheclyglvalhiisleueuSerleuCysergly 20
DB 147 ATGAGGCTCCCTTCCTTCCTGTCAT- - - - - GTTACCTTTAGGCTGGGCTTGGT 200
QY 21 LysalalleCylysaanglyleSerlysaRthphegiugliuleysglugliule 40
DB 201 AAACCTGATTCAGAAATGGGTTTCAGAGAACATTCGAGAAATTAAGAAATATA 260
QY 41 AlaseerCyaglaspvalalalyalalleleasneulaavalltyrglylysalagln 60
DB 261 GCCAAGCTATGAAAGATGTTGCTAAAGCATATACACCTTGCTTTATGTAATACAG 320
QY 61 AsnarSerlyTguatrgleualaleuuevalaspThrValGlyProatrgleuSerGly 80
DB 321 AACCGTCTATGAGCGCTTGGGACTTCTAGTTATCTGTTGGACCACTGAGTGGC 380
QY 81 SerlysaenleugliulysalallelelnlemerTyrglnaenleuglnlnaspglyleu 100
DB 381 TCTAAGAACCTAGGAAGAGCTATTCATATCATGACCAAACTCGACAGAGATGGCTA 440
QY 101 GlulysValhiisleugluProvalargileprohiistrpgluarglyglugluserala 120
DB 441 GAATATGTTACCTGGAGCAGGTGAGATACCCACTGGAGAGGAGAAATCTGCA 500
QY 121 ValmetleugluProatrglihehlylelyleallleleuglyleuglyserSerilegly 140
DB 501 GTGATGCTTGAGCCTCCAAATTCACAAAGATGGCTATTCTAGGCTTGGACAGCATGGG 560
QY 141 ThrproglugliyllietrAlagluValleuValValThrserepaspgluleugln 160
DB 561 ACTCCCGAGAGGACATCACAGAGAGTGTGTGTAGCTTTTGAATGAATCTCA 620
QY 161 ArgarglaserGluualarGlylyleValValTyraenglnProTyrlleasnTyrr 180
DB 621 AGAAGACATCAGAGAGAGAGGAGATCATCTTTATACACGCTTACACGCGCTAT 680
QY 181 SerargThrValGlnTyrraGthrGlnGlyAlaValaGluAlaAlaValaGlyAlaLeu 200
DB 681 GAGAAAGACTGTCAGTACCGGCTGCAGAGAGCTGTGGAAGCTGCCAAGTGGAGCTGTG 740
QY 201 AlaseerleuilearGserValAlaseerPheSerileTyserProhistrnglylleGln 220
DB 741 GCATCCCTCATCATCATCATGATGCTCTTTTCATCTACAGTCTCTCACACAGGATTCAG 800
QY 221 GlulTyrglnaspGlyValProlyleleProThraCyseileThrValGluaspAlaGlu 240
DB 801 AAATATCAAGATGCTGTGCCCAAGATTCCAACGCTGTATCACAGTAAAGATGCAGAA 860
QY 241 MetwetseryargwetalaserhiaglylelylevalleleugleulysmetGlyAla 260
DB 861 ATGATGCTCGAATGCGCTTCGCGGAGAAATAATTCATTCTCGAAATGGAGCA 920
QY 261 LysrhtTyrrProaspThrarpSerPheasnthValaGluileThrnglyserlyTyrr 280
DB 921 AAGCCATCATCAATCATCTCTTCAATACATCTTGCAAGATCATCTGGAGCATGTAT 980
QY 281 ProgluglnValalleuValaserGlyhiisleuaspSerTrpaspValGlynglnGlyAla 300
DB 981 CCAAGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1040
QY 301 MetaspaspGlygllyalaphelleSerTyrglnuAlaenSerleuilelyaspLeu 320
DB 1041 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1100

QY 321 GlyLeuArgProIysArgThrLeuArgLeuValLeuThrAlaIleGluGlnGly 340
 DB 1101 GGGCTGGCTCCAAAGAGAGACTCTGGCTGGCTGGCTGAGTGCAGAAAGAGAGGA 1160
 QY 341 ValGlyAlaPheGlnIleThrGlnLeuHisValAsnIleSerAsnIleThrLeuVal 360
 DB 1161 ATTGGTCTCCCTCCGCTATTAGAGCTACATAGCAATATATTTCCAGTACAGCTGGTG 1220
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlnLys 380
 DB 1221 ATGGACCTGACTGAGAGAACTCTTACCCACTGACTGAGCTGCTGCTGAGGAGAGG 1280
 QY 381 AlaArgAlaIleMetGluGlnValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 DB 1281 GCCAGGGCTATCATGAAAGAAAGTATGATCTTGGCAACCCCTCATGTCACCAAGTTC 1340
 QY 401 LeuSerHisGlyGlyGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyVal 420
 DB 1341 TTTAGTATGAGAGAGAACTGATATTAACCTTGGATCCAGCTGAGTACCTGAGACC 1400
 QY 421 SerLeuLeuAspAspLeuIleThrLeuThrPhePhePheHisSerHisGlyAspThrMet 440
 DB 1401 AGCTGGCTGATGACTGCTGACAGATATTTCTTTCCATCTCCACGAGAGACCAATG 1460
 QY 441 ThrValMetAspProIysGlnMetAsnValAlaAlaIleValAlaValSerTyr 460
 DB 1461 ACTGTCATGATCCAAAGCAGATGAATGTTCTGCTGCTGATGGCTGTGCTTAT 1520
 QY 461 ValValAlaAspMetGluGlnMetLeuProArgSer 472
 DB 1521 GTTGTTCAGACATGATGAATGCTGCCAGGTCC 1556
 RESULT 3
 LOCUS BC010977 1868 bp mRNA linear HTC 04-MAR-2003
 DEFINITION Mus musculus, similar to plasma glutamate carboxypeptidase, clone
 IMAGE:4018296, mRNA.
 ACCESSION BC010977
 VERSION BC010977.1 GI:15012137
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1868)
 REFERENCES
 AUTHORS Straubeberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabds-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Institute for Systems Biology
 http://www.systemsbio.org
 contact: amadan@systemsbio.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whitting
 Cloning distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
 Series: IRAX Plate: 18 Row: 5 Column: 4
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 9055233
 This clone has the following problem: retained intron.
 FEATURES
 SOURCE
 1. 1868
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"

US-09-745-763-36 (1-472) x BC010977 (1-1868)
 Alignment Scores:
 Pred. No.: 1,79e-234 Length: 1868
 Score: 2124.00 Matches: 410
 Percent Similarity: 94.79% Conservative: 27
 Best Local Similarity: 88.94% Mismatches: 24
 Query Match: 87.88% Indels: 0
 DB: Gaps: 0
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4018296"
 /tissue_type="Mammary tumor metastasized to lung. Tumor
 arose spontaneously from a senescent normal mammary
 (clonal) outgrowth infected with the virus MMV."
 /clone_id="NCI CGAP_Lu29"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"
 BASE COUNT 505 a 418 c 460 g 485 t
 ORIGIN
 QY 12 ValHisLeuLeuSerLeuCysSerGlyLeuAlaIleCysAlaSerGlyLeuValArg 31
 DB GTTCACTTATGAGCTTGGCTGCTGTAAGCTGTAAGCTGTAAGCTGTTCTCAGCA 287
 QY 32 ThrPheGlnIleValLeuGlyGlnIleAlaSerCysGlyAspValAlaValAlaIle 51
 DB AACATTCGAAATATTAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 347
 QY 52 AsnLeuAlaValIleGlyLysAlaGlnAsnArgSerTyrGlnArgLeuAlaLeuVal 71
 DB AACCTGCTGTTATGATGTAATACCAAGACCGGTCTTATGAGCTTCTGAGCTTCTG 407
 QY 348 AspThrValGlyProArgLeuSerGlySerLysAsnLeuGlnLysAlaIleGlnIleMet 91
 DB GATGCTGTGAGCCAGACAGTGAAGCTTGAAGAACTGAGAAAGCTTCAAAATCATG 467
 QY 408 TyrGlnAsnLeuGlnIleAspGlyLeuGlnValHisLeuGlnProValArgIlePro 111
 DB TACCAAACTGACGAAAGATGGCTGAAGAAATGTTCACTGAGCAGCTGAGAAATGCC 527
 QY 468 HisTrpGlnArgGlyGlnGlnSerAlaValMetLeuGlnProArgIleHisLysIleAla 131
 DB CACTGGAGAGGGGAGAGAAATCTGCACTGATGCTTGAAGCTTCAAGTTCACAGATGCT 587
 QY 528 IleLeuGlyLeuGlySerSerIleGlyThrProProGlnGlyIleThrAlaGlnValLeu 151
 DB ATTCTAGCTCTTGGCAGCAGCATTTGGGATCCCCAGAGGAGCATCACAGAGAAAGTCTG 647
 QY 132 ValValThrSerPheAspGluLeuGlnArgArgAlaSerGlnAlaArgGlyValIleVal 171
 DB GTGGTACCTCTTTTGTGATGAACTTCAGAAAGAGCATCAAGAGCAAGAGGAAATCAT 707
 QY 648 ValTyrAsnGlnProTyrIleAsnTyrSerArgThrValGlnTyrArgThrGlnValAla 191
 DB GTTTATACCAAGCTTTCACCTGCTTGAAGAGCTGCGATGCCGGTGCAGAGAGCT 767
 QY 708 ValGlnAlaAlaLysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSer 211
 DB GTGGAAGCTGCAAGAGTGGAGGCTGGGATCCCTCATCAATCAGATGCTCTTCTTCC 827
 QY 768 IleTyrSerProHisThrGlyIleGlnGlnTyrGlnAspGlyValProValIleProThr 231
 DB ATTCACAGTCTTCACAGAGTATTCAGAAATATCAAGATGCTGTCAGCAAGATTCACA 887
 QY 828 AlaCysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGlyIleLys 251
 DB GCTGTTATCAAGTGAAGATGCAAGAAATATGTTGTAATGGCTTCTGTCGGAACAAA 947
 QY 888 IleValIleGlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPheAsnThr 271
 DB ATTGTCAATTCATCTGAAATGGGAGCAAGACCTATCCAGATCTGATTCCTTCAATACT 1007

Oy 272 ValAlaGluIleThrGlySerIleTyrProGluGlnValIleuValSerGlyHisIleu 291
 Db 1008 GTTGAGAGATCAGTCCGAGAGCATGTATCCAGAGAAAGTTCTCTGTCAGTGCAGCATTTG 1067
 Oy 292 AspSerTyrAspValGlyGlnGlyValIleuValAspGlyGlyValAlaPheIleSerTyr 311
 Db 1068 GACAGCTGGAGATGTTGGGAGAGGTCCTCGATGATGGTGGTGGAGCCCTTATCATCAGG 1127
 Oy 312 GluAlaLeuSerIleIleValAspIleuValLeuArgProValArgThrIleuArgIleuVal 331
 Db 1128 GAAGACCTCTCACTTGTAAATATCTTGGGCTGCGCTCCAAAGAGACCTTGGCGCTGGTG 1187
 Oy 332 LeuTyrThrAlaGluGlnGlyGlnGlyValGlyValAlaPheGlnTyrGlnIleuHisIle 351
 Db 1188 CTCGACCTGCGAGAGAAACAGAGAGATGTGTCTCTCCAGTATATGAGTACATTAAG 1247
 Oy 352 ValAlaIleSerAspTyrSerIleuValMetGluSerAspAlaGlyThrPheLeuProThr 371
 Db 1248 GCAATATATTTCCAGTACAGCTTGGTGGATGAGAAAGCTGACAGAACCTTCTTACCACT 1307
 Oy 372 GlyLeuGlnPheThrGlySerGlyValAlaArgAlaIleMetGluGlnValMetSerIleu 391
 Db 1308 GGAGTCGAGTTCACGCGAGAGTGAAGCCAGGCTATCATGAGAGAACTATGATCTT 1367
 Oy 392 LeuGlnProLeuAspIleThrGlnValIleuSerHisGlyGlyGlyThrAspIleAspPhe 411
 Db 1368 CTGCAACCCCTCAATGTCACCAAGTCTTTAGTATGAGAGAGAAAGTATTTACTTC 1427
 Oy 412 TrrIleGlnAlaGlyValProGlyValAspIleuLeuAspAspLeuTyrIleTyrPhePhe 431
 Db 1428 TGGATCCAGAGTGAAGTACCTGAGCCAGTCTGCGATGATGATGAGAAAGTATTTCTTT 1487
 Oy 432 PheHisIleSerHisGlyAspThrMetThrValMetAspProGlyGlnMetAsnValAla 451
 Db 1488 TTCCATTCATTCCTCAGAGAGACCAATGATGATGATGATGATGATGATGATGATGATG 1547
 Oy 452 AlaAlaValTrrAlaValAlaValSerTyrValValAlaAspMetGluGlnMetLeuProArg 471
 Db 1548 GCTGCTGTGTGGCTGT 1607
 Oy 472 Ser 472
 Db 1608 TCC 1610
 RESULT 4
 BC012019 1024 bp mRNA linear HTC 04-MAR-2003
 LOCUS Homo sapiens, similar to plasma glutamate carboxypeptidase, clone
 DEFINITION IMAGE:4455631, mRNA.
 ACCESSION BC012019.1 GI:15080558
 VERSION 1
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1024)
 Strausberg, R.
 DIRECT SUBMISSION
 TITLE Submitted (30-JUN-2001) National Institutes of Health, Mammalian
 JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanaavali,
 A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
 Series: IRAC Plate: 28 Row: B Column: 11
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 5174626
 This clone has the following problem: retained intron.
 Location/Qualifiers
 FEATURES
 source
 1. 1024
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4455631"
 /tissue_type="Kidney, hypernephroma"
 /clone_id="NIH MGC_89"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"
 BASE COUNT 317 a 210 c 256 g 241 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,23e-178 Length: 1024
 Score: 1633.00 Matches: 323
 Percent Similarity: 99.09% Conserved: 2
 Best Local Similarity: 98.48% Mismatches: 1
 Query Match: 67.56% Indels: 2
 DB: 11 Gaps: 1
 US-09-745-763-36 (1-472) x BC012019 (1-1024)
 Oy 1 MetIysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20
 Db 36 ATGAATTCCTTATCTTCGACATTTTGGTGGTGTCACTTTATCTCTGCTGCTGGG 95
 Oy 21 LysAlaIleCyIysAsnGlyIleSerIysArgThrPheGluGluIleValIle 40
 Db 96 AAAGCTATATGCAAGATGCGATCTTAAGAGACTTTTGAAGAAATTAAGAAATTA 155
 Oy 41 AlaserCyGlyAspValAlaIleValAlaIleIleAsnLeuAlaValTyrGlyValAlaGln 60
 Db 156 GCCAGCTGTGAGATGTTGCTTAAGCAATCACTACCTGCTTTATGTAAGCCAG 215
 Oy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
 Db 216 AACAGATCTTATGAGCATTTGCGACTTCTGTGATATCTGTTGAGACCAAGTGAAGTGC 275
 Oy 81 SerIysAsnLeuGluValAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
 Db 276 TCCAGAGACCTTAAGAAAAAGCCATCAATTAATGTAACCAAACTGCGAGCAAGATGGCTG 335
 Oy 101 GluIysValHisLeuGluProValArgIleProHisTrrGluArgGlyGlnGlySerIleGly 120
 Db 336 GAGAAAGTTCACTGAGCCAGTGAATACCCACCTGCGAGAGAGAGAGATCAAGTCACT 395
 Oy 121 ValMetLeuGluProArgIleHisIysAlaIleIleLeuGlyLeuGlySerIleGly 140
 Db 396 GTGATGCTGAGCAAGAAATCATTAAGTACCACTCTGCGGTCTTGGCAGCAGCATTTGGG 455
 Oy 141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
 Db 456 ACTCTCTCAGAGGACATTACAGCAAGAGTTCTGTGTGAGTCACTTTGAGAACTGAG 515
 Oy 161 ArgArgAlaSerGluAlaArgGlyValIleValValTyrAsnGlnProTyrIleAsnTyr 180
 Db 516 AGAAGGGCTTGAAGAGCAAGAGAGAGTGTGTGTTATTAACCAACTTATCAATCACTAC 575
 Oy 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaAlaIysValGlyAlaLeu 200

Db	576	TCAAGGACGGTGCATATCCGAACGAGGCGGCTGGAAAGTGCAAGTTGGGCTTTG	633
Oy	201	Alaserleu1lea9serValAlasertheser1leySerProHsrThGlytleIn	220
Db	636	GCACTTCCTCATTCATCCGTGGCCTCTTCCTTCATCTCACGTCCCTCACAGATTTCAG	695
Oy	221	GIUtyrGINApgYvalProllysileProthrAlaCySilethrValGluaspAlaGlu	240
Db	696	GAATACCAGAGATGGGTGCCCAAGAATTCACAACGCTGTATTACGGTAGGAAGATGCGAA	755
Oy	241	MetweseArAmetalaserHisgIyleysilleVallleGlnLeuLyMetGlyVala	260
Db	756	ATGATGTCAGAAATAGGCTTCTTCATCGAGATCAAANTGTCACTTACGCTPAAGATGGCGCA	815
Oy	261	LysThrTyPrOaApThrAspSerPheanThrValAlaGluilethrGlySerlyStyr	280
Db	816	AAGACTPACCCAGATACTGATTCCTTCACACTGACAGAGATCACTGGAGCAAAATAT	875
Oy	281	ProGUINValVallleuValSerGlyHisleuaspserTPAAPValGlyGlnGlyVala	300
Db	876	CCAAGAACGGTTGATCGGTGACAGAGACATCGACACACTGGAGTGGGCGAGGTC	935
Oy	301	MetAspAPgJylGlyGlyAlaphelieserTrpGuaAlaeusertleu1ltyAspleu	320
Db	936	ATGATGATGTCGGCGGTGGAGCCTTTATATACAGGAGACCTCTCACTTAATAAGATCTT	995
Oy	321	Gly-----LeuarGProlyArg 326	
Db	996	GGTAAATATTTAAAAAAAATAA 1019	
RESULT 5	BX439467	1050 bp mRNA linear EST 15-MAY-2003	
LOCUS	BX439467	Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE010YJ06	
DEFINITION	BX439467	5-PRIME, mRNA sequence.	
ACCESSION	BX439467		
VERSION	BX439467.1	GI:30777746	
KEYWORDS	EST.		
SOURCE	Homo sapiens	(human)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1050)		
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	Library was constructed by life technologies, a division of		
	Invitrogen. This sequence belongs to sequence cluster 4663.f For		
	more information about this cluster, see		
	http://www.genoscope.cns.fr/cgl-bin/cluster.cgi?seq=CS0DE010DEB03QPlc&cluster=4663.f. Contact :		
	Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/invitrogen Corporation 1600		
	Faraday Avenue Genoscope sequence ID : CS0DE010DEB03QPl.		
FEATURES	location/Qualifiers		
source	1..1050		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0DE010YJ06"		
	/issue_type="PLACENTA"		
	/clone_lib="Homo sapiens PLACENTA"		
	/note="Vector: pCMWSPORT.6; 1st strand cDNA was primed		
	with a NotI-oligo(drr) primer. Five prime end enriched,		
	double-strand cDNA was digested with Not I and cloned into		
	the Not I and EcoRV sites of the pCMWSPORT 6 vector.		
	Library was not normalized."		
BASE COUNT	307 a 233 C 271 G 235 T		
ORIGIN			

Alignment Scores:		Pred. No.:	4.1e-167	Length:	1050
Score:	1539.00			Matches:	303
Percent Similarity:	100.00%			Conservative:	0
Best Local Similarity:	100.00%			Mismatches:	0
Query Match:	63.67%			Indels:	0
DB:	13			Gaps:	0
US-09-745-763-36 (1-472) x BK439467 (1-1050)					
QY	1	MetIysPheLeuIlePheAlaPheBheGlyGlyValHisLeuLeuSerLeuCySergly	20		
DB	142	ATGAATTCCTTATCTTCGATTTTTCGGTGGTGTTCACCTTTATCCCTGCTCTGGG	201		
QY	21	LysAlaIleCyeIlySaengIyIleSerIysaRgthrPheGluGluIleIySegIuGluIle	40		
DB	202	AAAGCTAATTCAGAAAGATGGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAA	261		
QY	41	AlaSerCySgIyAspValAlaIlyAlaIleIleAenLeuAlaValIyGlyIySAlaGln	60		
DB	262	GCCAGCTGTGAGATGTTGCTTAACCATTCATCAACCTGCTTTATGTGTAAGCCG	321		
QY	61	AsnArgSerIyrgIuArgLeuAlaLeuLeuValaAspThrValGlyProArgLeuSergly	80		
DB	322	AACAGATCCTTAAGACGATGGCATCTTCTGTTGTATCTGTTGTGACCCAGACTGATGGC	381		
QY	81	SerIysaenLeuGluValaIleGlnIleMetIyrgIlnaenLeuGlnIlnaSpIyLeu	100		
DB	382	TTCCAGAAACCTTAAGAAAGGATCCAAATTATGTACCAAACTGACAGCAAGATGGCTG	441		
QY	101	GluIysValHisLeuGluProValArgIleProHisSTPGIuArgGlyGluGluSerAla	120		
DB	442	GAGAAAGTTCACTGAGACCGATGAGATTAACCCACTGGGAGAGGGAGAGAAATCAGCT	501		
QY	121	ValMetLeuGluProArgIleHisIySAlaIleAlaIleuGlyLeuGlySerSerIleGly	140		
DB	502	GTGATGCTGGAGCCAAAGATTCATTAAGATGACCATCTGGGCTTGGCAGCAGCTTGGG	561		
QY	141	ThrProProGluGlyIleThrAlaGluValLeuValIValThrSerPheAspGluLeuGln	160		
DB	562	ACTCTTCAGAAAGGATTAAGCAAGCAAAATTCGTGGTGAGACCTCTTTCGATGAACCTGC	621		
QY	161	ArgArgAlaSerGluAlaArgGlyIySAlleValIyTrAsnGlnProTyrIleAsnTyr	180		
DB	622	AGAAAGGCTCAGAGCAAGAGGAGAGATGTTGTTTATTAACCAACTTACATCACTAC	681		
QY	181	SerArgThrValGlnTyrArgThrGlnIyAlaValaGluAlaAlaIySValaGlyAlaLeu	200		
DB	682	TCAAGAGCGGTGCAATTCAGAAACGAGGGGGGGCGTGAAGCTGCCAAGGTTGGGCTTGG	741		
QY	201	AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrIyIleGln	220		
DB	742	GCATTCCTCATTCGATCCGTGGCTCTCTTCATCTCAAGTCCACAGCTCCACACAGATATTCG	801		
QY	221	GluTyrGlnaSpGlyValProIySleProThrAlaCySAlleThrValaGluAspAlaGlu	240		
DB	802	GAAATCCAGATGGGCTGCCCAAGATTCACACAGCTGTATTAACGTTGGAAGATCAGAA	861		
QY	241	MetMetSerArgMetAlaSerHisGlyIleIySAlleValaIleGlnLeuIySmetGlyAla	260		
DB	862	ATGATGTCAAGATAGCTTTCATGTGGATCAAAATTGTCATTCACACTAAAGATGGGGCA	921		
QY	261	LysThrTyrProAspThrAspSerPheAsnThrValaIaGluIleThrGlySerIySyr	280		
DB	922	AAGACTTACCCAGATACGTATTCCTTCAACACTGTAGCAAGATACCTGGGAGCAATAT	981		
QY	281	ProGluGlnValaIleuValSerGlyHisLeuAspSerTyrAspValaGlyGlnIyAla	300		
DB	982	CCAGAACAGGTTGTACTGTGATGATGATCTGACATCTGACAGCTGGGATGTTGGGAGGTTGC	1041		
QY	301	MetAspAsp	303		

```

Db      1042 ATGATGAT 1050

RESULT 6
LOCUS   BX355940
DEFINITION BX355940 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0D1006Y115 5-PRIME, mRNA sequence.
ACCESSION BX355940
VERSION  BX355940.1 GI:30384019
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1038)
AUTHORS  Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE    Full-length cDNA libraries and normalization.
JOURNAL  Unpublished
COMMENT  Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f for more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1006AE08QPLcluster=4663.f. Contact : Feng Liang Email : fliang@life.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1006AE08QPL.

FEATURES
source
1..1038
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1006Y115"
/tissue="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 304 a ~237 c 262 g 233 t
ORIGIN
Alignment Scores:
Pred. No.: 1,28e-154 Length: 1038
Score: 1431.00 Matches: 286
Percent Similarity: 99.65% Conservative: 0
Best Local Similarity: 99.65% Mismatches: 0
Query Match: 59.21% Indels: 1
DB: 13 Gaps: 0

US-09-745-763-36 (1-472) x BX355940 (1-1038)
Oy      1 MetLysPheLeuIiPhehAlaPhehGlyValHisLeuLeuSerLeuCySerGly 20
Db      186 ATGAATTCCTATCTTCGACATTTTCGGTGTGTTACCTTTTACCTGCTCTGG 239
Oy      21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPhehGluIleLysGluIle 40
Db      240 AAAGCTATATGCAAGATGCGATCTCTAAGAGCATTTTGAAGAATAAAGAAATA 299
Oy      41 AlaSerCyGlyAspValAlaLysAlaIleIleLeuLeuAlaValTyrGlyValAgl 60
Db      300 GCCACCTGTGGAGATGTCTAAACCAATCAACCACTAGCTTTATGTTAAACCCAG 359
Oy      61 AsnArgSerTyrGlyArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80
Db      360 AACCAATCTTATGACGATGCGACTTCTGTTGATACCTGTTGGACCCAGACTGAGTGC 419
Oy      81 SerLysAsnLeuGluLysAlaIleGluIleMetTyrGlnAsnLeuGlnAspGlyLeu 100

```

```

Db      420 TCCAAGAACTAGAAAAAGCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGCTG 479
Oy      101 GluLysValHisLeuGluProValArgIleProHisThrPgluArgGlyGluLysSerAla 120
Db      480 GAGAAAGTTCACCTGGAGCCAGTGAAGATACCCACCTGGAGAGGGAGAAATACAGCT 539
Oy      121 ValMetLeuGluProValGlyLeuHisIleAlaIleLeuGlyLysSerSerIleGly 140
Db      540 GTGATGCTGAGCCAGAAATTCATTAAGATAGCCATCTCGGTCTTGGACAGCATGGG 599
Oy      141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
Db      600 ACTCTCCAGAAAGCATTACAGCAAGATTCTGTGTGTGACCTCTTGATGAATCTCAG 659
Oy      161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
Db      660 AGAAGGCTCTAGAAAGCAAGGAGAGATGTGTGTTATTAACCACTTACATCACTAC 719
Oy      181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaValAlaValGlyValLeu 200
Db      720 TCAGAGAGGTGCAATACCAAGCAGGAGGCGGTGGAGCTGCAGAGTTGGGCTTTG 779
Oy      201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db      780 GCATCTCTCATTCGATCCGTGGCTCTCTCCATCTACAGTCCCTCACACAGATTTACAG 839
Oy      221 GluTyrGlnAspGlyValProValIleProThrValCysIleThrValGluAspAlaGlu 240
Db      840 GAATACCAAGATGAGCGGCCCAAGATTCACACACCTGATTAAGGAGGAGAGCAAGA 899
Oy      241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMet-GlyAla 260
Db      900 ATGATGTCAGAAATGGCTTTCATGGATCAAAATGTGCATTCAGTCAAGATGAGGGGCG 959
Oy      260 AluGlyThrTyrProAspPhePheSerPheAsnThrValAlaGluIleThrGlySerIle 280
Db      960 AAACACCTACCCAGATACGTATTCCTTCAACACATGTAAGATCACTGGAGCAATA 1019
Oy      280 rProGluGlnValValLeu 286
Db      1020 TCACGAACAGGTGTACTG 1038

RESULT 7
LOCUS   BO878966
DEFINITION AGENCOURT_8183107 Lupski_dorsal_root_ganglion Homo sapiens CDNA clone IMAGE:6184129 5', mRNA sequence.
ACCESSION BO878966
VERSION  BO878966.1 GI:22270974
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS  NIH-MGC http://mgs.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished
COMMENT  Contact: Robert Strausberg, Ph.D.
Email: cgabs@remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LRAM3572 row: n column: 02
High quality sequence stop: 696.
FEATURES
source
1..909
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:6184129"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_id="lupski_dorsal_root_ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGCTCG-3' and
5'-GACTAGTTCTAGATCGGAGCGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT      273 a      202 c      226 g      207 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      1,48e-145      Length:      909
Score:          1352.00      Matches:      276
Percent Similarity: 97.89%      Conservative: 2
Best Local Similarity: 97.18%      Mismatches: 3
Query Match:    55.94%      Indels:      3
DB:             13      Gaps:        0
US-09-745-763-36 (1-472) x BQ878966 (1-909)
QY      1 MetlyspheleuilephealaphhegilygllyvalhisleuenseuileuCyseerily 20
Db      64 ATGAATTCCTTATCTTCGATTTTTCGGTGTGTCACCTTTATCCCTGCTGCGG 123
QY      21 LyalaalleCylyspasnglylleserlyarhthreghugllylleeglugllyle 40
Db      124 AAACCTATATCCAGAAATGCGATCTCTTAAGAGGACTTTTGAAGAAATTAAGAAATA 183
QY      41 AlaserCysegiyarpvalalalyalalelleleaneuavalalyryglyvalagln 60
Db      184 GCCAGCTGTGAGATGTTGTCAAGCAATCAACCACTGATTATGTAAGAGCCAG 243
QY      61 Asnarserlyrygluarleualaleuvalasprhvalgllyproargleuserily 80
Db      244 AACGATCTTATGACGATTCGCACTTGTGTTATCTTGACCCAGACTGATGCG 303
QY      81 Serlyshanleuglylyvalallelellemetryrginaenleuglnaspolyleu 100
Db      304 TCCAAGAACTTAGAAAAGCCATCCAAATTAATGTAACAAAACCTGACAGATGGCTA 363
QY      101 Glulyrvalahisleugluprovalarglleprohistrpgluarglygluserila 120
Db      364 GAGAAAGTTCACTGAGACCACTGAGATACCCCACTGGAGAGGGGAGAAATCACT 423
QY      121 Valmetleugluproargllehislyrillealalleleuglyleuglyserserillegly 140
Db      424 GTGATCTCGAGCCAGAAATTCATTAAGTACCACTCTCGGCTTTGGACAGCATTTG 483
QY      141 ThrPrologllyllylethralagluvalleuvalalthrserpheaspolyleu 160
Db      484 ACTCCTCCAGAAAGGATTAACAGCAAGTTCTGGGTGGAGACTCTTGAGTAAGTCCAG 543
QY      161 Argarglasergluvalarglylyvalilevalialyrraenglnprotyrilleasntyr 180
Db      544 AGAAGGGGCTCAGAAAGAGGAGGAGATGTTGTTTAAACCAACTTAATCAACATC 603
QY      181 Serargthralvalglntyrargthrglnlyalvalagluvalalalyvalaglyalaleu 200
Db      604 TCAAGAGGCTGCATTAACCAAGCGAGGGGGCGGTGAGAGCTGCCAAGCTTGGGCTTG 663
QY      201 Alaserleuileargservalalaserpheserilletyrserprohistrhryllegln 220
Db      664 GCATCTCTCATTCGATCCGTGGCTCTTCTCACTCACTCACTCACTCACTCACTCACTCA 723

```

```

QY      221 GlutyrglnaspolyvalprolylleprothralaCysillethrrvalgluaspaaglu 240
Db      724 GAATACAGATGCGCGGCCCAAGAAATTCACACGCCCTGTTATTCGGTGGAGATGCAGAA 783
QY      241 Metwetserrhmetalaaserhlyllylevalilleleuauysmet-glyal 260
Db      784 ATGATGTCAGAAATGCGCTTCATGAGGATCAAAATTCATTCAGTAAAGATGGGGGC 843
QY      260 alvsthrtyrproasphrthrserspheantthr-valalagluillethr-glyserlyls 279
Db      844 AAACCTACCCAGATTAAGATTCCTTTCACACTGCTAGAGATCACTGGGAGCAAA 903
QY      280 Tyepro 281
Db      904 TATCCC 909
RESULT 8
BX360507
LOCUS
DEFINITION
BX360507 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
ACCESSION
BX360507
VERSION
BX360507.1 GI:30374439
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1126)
AUTHORS
Ll,w.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1071AA10Qp1cluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1071AA10Qp1.
FEATURES
Location/Qualifiers
source
1..1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1071AA19"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      311 a      249 c      298 g      246 t      22 others
ORIGIN
Alignment Scores:
Pred. No.:      7.43e-141      Length:      1126
Score:          1312.50      Matches:      282
Percent Similarity: 94.02%      Conservative: 1
Best Local Similarity: 93.69%      Mismatches: 14
Query Match:    54.30%      Indels:      8
DB:             13      Gaps:        2
US-09-745-763-36 (1-472) x BX360507 (1-1126)
QY      1 MetlyspheleuilephealaphhegilygllyvalhisleuenseuileuCyseerily 20
Db      210 ATGAATTCCTTATCTTCGATTTTTCGGTGTGTCACCTTTATCCCTGCTGCGG 269

```

Oy 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLeuGluIle 40
 Db 270 AAAGCTATATGCAAGATGGCATCTCTAAGAGACTTTTGAAGAAATTAAGAAATA 329
 Oy 41 AAserCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
 Db 330 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 389
 Oy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
 Db 390 AACAGATCTTAGAGAGATGGCATCTCTGTTGATGTTGAGCCAGATGTGATG 449
 Oy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100
 Db 450 TCCAGAACTTAGAAGAAAGCCATCCAAATTAATGTCACAAACCTGAGAGATGGCTG 509
 Oy 101 GluLysValIleLeuGluProValArgIleProIleStrpGluArgGlyGluGluSerAla 120
 Db 510 GAGAAAGTTTCACTGGAGCAGTGAAGATACCCCACTGGAGAGGGGAGAAATCAGCT 569
 Oy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
 Db 570 GTGATGCTGAGGCCAAGAAATTCATAAGATAGCATCTGGGCTTGGCAGCAGCATTTGG 629
 Oy 141 ThrProProGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160
 Db 630 ACTCTTCAGAAAGGCAATACAGCAAGATTTCTGGTGAGTCACTCTTTCATGATTAATGCG 689
 Oy 161 ArgArgAlaSerGluAlaArgGlyLysIleValIleTyrAsnGlnProTyrIleAsnTyr 180
 Db 690 AGAAGGCTCTGAGAGCAAGAGAGATGTTGTTTAAACCAACTTACATCACTAC 749
 Oy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200
 Db 750 TCAAGAGAGCTGTCATACCGAAGCGAGGGGCGGAGAGCTGCCAAGGTTGGGCTTTG 809
 Oy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
 Db 810 GCATCTCTCATTCGATTCCTGGCTCTCTTCTTCATCTACAGCTCCACAGATTTTCCG 869
 Oy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrVal-GluAspAlaGln 240
 Db 870 GAATACCAAGATGGCTGGCCCAAGATTCACAGCTGATATACCGTGGAGAGATGCA 929
 Oy 240 UMeMetSerArgMetAlaSerHis-GlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 Db 930 AATGATGTCAGAAATGGCTTCTCATTTGGGATCAAAATTTGTCACTCAAGTAAAGATGG 989
 Oy 260 ValLeuThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLys 280
 Db 990 S-AAAGACTACCCAGATACCTGATCTTTCAC-ACGTGACCAAA-ATCACTGGAGCAAA-W 1045
 Oy 280 TyrProGluGlnValValLeuValSerGlyHisLeuAspSerTyrPaspValGlyGlnGly 299
 Db 1046 ATCCAAAGAGCTKTACT---GKYAGTGGW---WCTGGRWASTGGATGTTGGWMMGGGG 1098

RESULT 9
 BG751497 902 bp mRNA linear EST 15-MAY-2001
 LOCUS 60230076f1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4873770 5',
 DEFINITION mRNA sequence.
 ACCESSION BG751497
 VERSION BG751497.1 GI:14062150
 KEYWORDS EST.
 ORGANISM Homo sapiens (human)
 SOURCE Eukaryote; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 902)
 AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished

COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: rs@amail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
 Plate: LICM1752 row: 5 column: 19
 High quality sequence stop: 817.
 Location/Qualifiers
 1..902
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4873770"
 /rname="IMAGE:4873770"
 /rsize="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 43"
 /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAAGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 246 a 200 c 251 g 204 t 1 othrs
 ORIGIN

Alignment Scores:
 Pred. No.: 1,02e-139 Length: 902
 Score: 1301.50 Matches: 279
 Percent Similarity: 93.77% Conservative: 7
 Best Local Similarity: 91.48% Mismatches: 6
 Query Match: 53.85% Indels: 16-
 DB: 10 Gaps: 3

US-09-745-763-36 (1-472) x BG751497 (1-902)
 Oy 111 ProHisTrpGluArgGlyGluSerAlaValMetLeuGluProArgIleHisLysIle 130
 Db 2 CCCCACTGGAGAGAGGGAGAAAGATCACTGTGATGCTGAGCCAAAGATTCATTAAGATA 61
 Oy 131 AlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIleThrAlaGluVal 150
 Db 62 GCCATCTCGGCTCTTGGCAGCAGATTTGGCACTCTCCAGAAAGCATTAACGAGAAATT 121
 Oy 151 LeuValValThrSerPheAspGluLeuGlnArgArgAlaSerGluAlaArgGlyLysIle 170
 Db 122 CTGGTGATGACCTTCTTCATGATCACTGCAGAGAGGGCTTCAGAGCAAGGAGAGAT- 180
 Oy 171 ValValTyrAsnGlnProTyrIleAsnTyrSerArgThrValGlnTyrArgThrGlnGly 190
 Db 181 GTTGTTATATACCAACCTTATCATCACTCAAGAGAGGTCAGATACCGAAGCAGGGG 240
 Oy 191 AlaValGluAlaAlaLysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPhe 210
 Db 241 GCGGTGAGAGCTGGCAAGGTGGGCTTTGGCATCTTCATTCATCGTGGCTCTTCC 300
 Oy 211 SerIleTyrSerProHisThrGlyIleGlnGluTyrGlnAspGlyValProLysIlePro 230
 Db 301 TCCATCTACACTCTCCACACAGATTCAGAAATTCAGAGATGGCGTCCCAAGATTTCCA 360
 Oy 231 ThrAlaCysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGlyIle 250
 Db 361 ACAAGCTGTATTAAGGTGAGAGATGCAAGATTAATGTCAAAGATGGCTTCAATGAGATC 420
 Oy 251 LysIleValIleGlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPheAsn 270
 Db 421 AAAATTGTCATTCAGCTAAAGATGGGCGCAAGAGCATCAACCAAGATTAATGATTCCTTCAAC 480

QY 271 ThrValAlaGluIleThrGlySerIleTyProGluGlnValIleValSerGlyHis 290
 DB 481 ACTGTAGACAGATCACTGGAGCAATATCCAGAACAGGTGTGCTGCTCAGTGACAT 540
 QY 291 LeuAspSerTTPAPPAValGlyGlnGlyAlaMet-AspAspGlyGlyGlyAlaPheIleSe 310
 DB 541 CTGACACAGCTGGAGTGGGAGGAGGCTCCATGCGATATGCGGTGAGGCTTTATATC 600
 QY 310 rTrGluValLeuSerLeuIleIleValAspLeuGlyLeuAspProGlyValGlyThrLeuArgIle 330
 DB 601 ATGGAGACACTCTCACTTATTAAGATCTTGCGCTGCTCCAGAGAGACTTGCGGCT 660
 QY 330 uValIleuTrpThr-AlaGluGluGlnGlyGlyValGlyAlaPheGlnTyTrpGlnLeuH 350
 DB 661 GGTGCTCTGACCTGGCAGAACAGAGGTGAGTGGTGGTCCCTCCAGATTTATGAGTTAC 720
 QY 350 IseIleValAsnIleSerAsn-----TyrSerIleuValMetGluSerAspA 365
 DB 721 ACAAGGTAAATATTTCCAACTAACAGTCTGGGTGATGAGATCTT-----GAGC 768
 QY 365 IAGIYThrPheLeu---ProThrGlyLeuGlnPheThr-GlySerGluValAlaArgAla 383
 DB 769 CAGAACCTTTCTTACCCACTGGGCTGCAATTCACCTTGGCAGTGAAGAGCCAGGCC 828
 QY 384 IleMetGluGluValMetSerIleuGlnProLeuAsnIleThrGlnValLeuSerHis 403
 DB 829 ATCATGAGAGAGGTTATGAGCTGATGAAGCCT--CATATCAGTCAAGTCCG-AAACAT 885
 QY 404 GlyGluGly 406
 DB 886 GGAAGAGGA 894

RESULT 10

BX335995

LOCUS BX335995 919 bp mRNA linear EST 02-MAY-2003
 DEFINITION BX335995 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
 clone CSOD1022YJ16 5-PRIME, mRNA sequence.

ACCESSION

BX335995

VERSION BX335995.1 GI:30339459
 KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4663.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSOD1022DE080P1&cluster=4663.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CSOD1022DE080P1.

FEATURES

source

1..919

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSOD1022YJ16"

/issue_type="PLACENTA COT 25-NORMALIZED"

/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT

273 a

206 c

226 g

208 t

ORIGIN

Alignment Scores:

Pred. No.: 2.25e-138 Length: 919
 Score: 1290.00 Matches: 262
 Percent Similarity: 98.87% Conservative: 0
 Best Local Similarity: 98.87% Mismatches: 2
 Query Match: 53.37% Indels: 2
 DB: 13 Gaps: 0

US-09-745-763-36 (1-472) x BX335995 (1-919)

QY 1 MetIysPheLeuIlePheAlaPheGlyGlyValHisLeuLeuSerLeuCyseSerGly 20
 DB 126 ATGAATTCCTTATCTTGCGATTTTGGGTGATGTTTACCTTTATCCCTGTGCTGGG 185
 QY 21 LysAlaIleCysIysAsnGlyIleSerIleArgThrPheGluGluIleLeuGluIle 40
 DB 186 AAACATATGCAAGAAATGGCATCTTAARAGCACTTTGAAAGAAATPAAARAGAAATA 245
 QY 41 AlAserCyseGlyAspValAlaIleValAlaIleIleAsnLeuAlaValTyrglyValAlaGln 60
 DB 246 GCCAGCTGTGGAGATGTTGCTAAAGCAATCAACCTAGCTGTTATGTAAAGCCAG 305
 QY 61 AsnArgSerTyrgIuArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80
 DB 306 AACGATCTTATGAGCGATTTGGCATCTTCTGTTATCTGTGAACCCAGACTAGTGC 365
 QY 81 SerIysAsnLeuGluValAlaIleGlnIleMetTyrgIuAsnLeuGlnIleAspGlyLeu 100
 DB 366 TCCAAGAACTTAAAGAAAGCCATCAATTTATGACCAAACTGCAGCAGAGATGGGCTG 425
 QY 101 GluIysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluSerAla 120
 DB 426 GAGAAAGTTCCTGAGGCCAGTGAATATCCCACTGGAGAGGAGGAGAAATACAGCT 485
 QY 121 ValMetLeuGluProValGlyIleHisIleValIleAlaIleLeuGlyLeuGlySerSerIleGly 140
 DB 486 GTGATGCTGAGGCCAAGATTCATTAAGATAGCCATCTCGGCTCTTGGCAGCAGCATTCGG 545
 QY 141 ThrProGluGlyIleThrAlaGluValIleuValIleThrSerPheAspGluLeuGln 160
 DB 546 ACTCTCCAGAAAGCATTAACAGCAGAAAGTTCTGTGCTGACCTCTTCAGATGAACTGCAG 605
 QY 161 ArgArgAlaSerGluAlaArgGlyIleValIleValIleValIleValIleValIleValIleVal 180
 DB 606 AGAAGGGCTCCAGAGAGAGAGGAGAGATTTGTTTATTAACCAACCTTATCATCACTAC 665
 QY 181 SerArgThrValGlnTyrgIuArgThrGlnGlyValAlaValGluAlaIleValIleValIleVal 200
 DB 666 TCAAGAGCGTGCATATCCAGAACCCAGGGGCGGTGAGAGCTGCCAAGTTGGGGCTTTG 725
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrgIuProHisTrpGlyIleGln 220
 DB 726 GCATCTCTCATTCAGTCGTCGTCCTCTTCATTAAGATAGCTCTTCACACAGGATTCAG 785
 QY 221 GluTyrgIuAspGlyValProValIleProThrAlaCysIleThrValIleGluAspAlaGlu 240
 DB 786 GAATTAACAGATGCGCGGCCAACAAATTCACACACCTGTATTAAGGTGAAGAGACAA 845
 QY 241 MetMetSerArgMetAlaSerHis-GlyIleIleValIleValIleGlnLeuIleValIleValIle 260
 DB 846 ATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTCTTCAAGCTTAARATGGGGGC 905
 QY 260 alyeThrTyrgPro 264
 DB 906 -AARACTACCA 917

RESULT 11

BUI46905

LOCUS

DEFINITION

AGENCOURT 7978204 Lupski dorsal root ganglion Homo sapiens CDNA

clone IMAGE:6185233 5', mRNA sequence.

ACCESSION	BUI146905
VERSION	BUI146905.1
KEYWORDS	GI:22660437
SOURCE	EST.
ORGANISM	Homo sapiens (human)
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
JOURNAL	1 (bases 1 to 866)
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Srausberg, Ph.D. Email: csgrabs-remail.nih.gov Tissue Procurement: Dr. James R. Lupski CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Consortium Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.jnl.gov Plate: LLM13575 row: 1 column: 02 High quality sequence stop: 731.
FEATURES	Location/Qualifiers
source	1..866 /organism="Homo sapiens" /mol_type="RNA" /db_xref="taxon:9606" /clone="IMAGE:6185233" /sex="male" /tissue_type="dorsal root ganglia" /dev_stage="adult, 36 yr" /lab_host="DH10B" /clone_lib="Lupski_dorsal_root_ganglion" /note=Vector: pCMV-SPORT6 (Life Technologies); Site:1: NotI; Site 2: SalI; CDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCGCAGCCCTCG-3' and 5'-GACTATGTTCNATGCCGACGGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylour College of Medicine) and is available through Life Technologies."
BASE COUNT	257 a 182 c 233 g 191 t 3 others
ORIGIN	
Alignment Scores:	
Pred. No.:	6.62e-137 Length: 866
Score:	1277.00 Matches: 268
Percent Similarity:	92.54% Conservative: 5
Best Local Similarity:	90.85% Mismatch: 11
Query Match:	52.83% Indels: 11
DQ:	13 Gaps: 4
DS-09-745-763-36 (1-472) x BUI146905 (1-866)	
OY	23 TLeCyLySaNgLyLiSeSerLySaRgThr-PheGngUuLiLyGsGUjUlLeAlase 42
Db	3 ATATGCAAGAATGGCATCTCTAAGAGAGCATTTNTGGAAGAAATAAAGAAATAGCCAG 62
OY	42 rCYgSLyASpVALALyALALyALALALellLLeLnemULaVALTYrgLYySAInAnAR 62
Db	63 CTGGGAGAGTGTCCTAAGAACATCAACTCAAGCTGTTTGGTAAAGCCCAAGACAG 122
OY	62 gSerTyGLArXLeJLaLeuLeuValAsPThrValGlYProAgLeuSerGlySerTy 82
Db	123 ATCTTAAGACGATTGGCACTTCTGATTGAATCTGTGGACCACAACAGTAGTGCTCAA 182
OY	82 sAsnLeugLUyLAALALegLinLeMeTYrgLnSnLeungIngaInAPgLyLeugLUy 102
Db	183 GAACCTGAAAAGCCATCCAATTATGTACCAAAACGACGACAAAGATGGGCTGGAGAA 242
OY	102 sValHLseugLUprOVAALrgLlePrOHsTrIpGLuArGLyguJUserALaValMe 122

Db	Accession	Definition	LOCUS	Result
QY	243	AGTTCACTGAGGACAGATGGAATATCCCACTGGGAGAGGAGGAAGATACAGTGTGAT	CB958693	302
QY	122	ctleugluproarglllehisylslllealaleleuglylsestertlleeglythrpr	CB958693	142
Db	303	GCTGGACCCAAAGATTCATTAAGATAGCCATCTGGGCTTTGGCGAGCATTTGGGACTCC	CB958693	362
QY	142	oprogunglylllsthprlagnlualleuvalvalthsetpheaspglueuglnaagat	CB958693	162
QY	363	ttcagaaaggcatrracagacaaagattctgggggggactcttttcataatgacgcagaaag	CB958693	422
QY	162	galaserglualaargyllysyllevalvaltyrasnglnprotyrlleantyrserar	CB958693	182
Db	423	ggcttcagaaagcaagaggaagantgtgtttatpaccaccttaacatcactcacaag	CB958693	482
QY	182	gthrvalglntyrargthrglnglvalavalglualalalysvalglvalaleuallase	CB958693	202
Db	483	gacggtgcaatpaccgaacgacgagggggcggtggagagctccaaagggtggcgcttggcattc	CB958693	542
QY	202	rleuillleargservalalaserpheserilletyrserprohissthrglyllelngluty	CB958693	222
Db	543	ttctatrrcgaatccgtggcctcctcttccatrracgctccacacagctatrrccagata	CB958693	602
QY	222	rglnaspglvalprouylsileprrhrralacylllethrvalgluabpallagluweme	CB958693	242
Db	603	ccagagagggcggtgcccaagattccaaacgctgattatccggtgabaatgcagaaatgat	CB958693	662
QY	242	tserargmetalasethisgllylleyslllevallleghleuulysmetglvalalysth	CB958693	262
Db	663	gtcagaaatggtcttctcatggagatcaaaattgtcattcacctaaagattggggcnaagac	CB958693	722
QY	262	rttyrproaspthraspserpheasnthrrvalagluillethrglyserlysttyrpro	CB958693	281
Db	723	ctaccgaatrrcgaatrrcctttnccactgtgagcagagatcactggagcaaatrrtcca	CB958693	782
QY	282	gluglnvalvalleuvalser---glyhisleuabp---sertrasp---valglgln	CB958693	298
Db	783	gaacaggggttgcactgggtcagtgcgacatctggaaacaancttggaaatgtttggggcaa	CB958693	842
QY	299	glyalawetaspaspglyglylalyalpheillestertp 311	CB958693	311
Db	843	ggggcgcattcgggca-----attgaattgg 866	CB958693	866

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30351744"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_184"
/note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site 1:
SfiI (ggccgcttcgagc); Site 2: SfiI (ggccgcttcgagc);
Library is oligo-dT primed and directionally cloned. cDNA
was prepared from a glandular pool of tissues from thyroid,
parathyroid, adrenal, cortex and pineal gland. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCGCATTTAGCC-3' and 3' adaptor sequence:
5'-ATCTAGAGCGCCGAGCGCCGACATG-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.38
kb (range 0.60-3.5 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

```

BASE COUNT 192 a 171 c 195 g 181 t

ORIGIN

Alignment Scores:

Pred. No.:	3,18e-134	Length:	739
Score:	1253.00	Matches:	244
Percent Similarity:	99.59%	Conservative:	0
Best Local Similarity:	99.59%	Mismatches:	0
Query Match:	51.84%	Indels:	1
DB:	14	Gaps:	0

US-09-745-763-36 (1-472) x CB958693 (1-739)

```

Qy 197 ValGlyAlaLeuAlaSerLeuLeuSerValAlaSerPheSerIleTyrSerProHis 216
Db 4 GTTGGGGGCTTGGCATCTCTCATTCGATCCGCGCTCTTCCATCTACAGTCTTAC 63
Qy 217 ThrGlyIleGlnGlyTyrGlnAspGlyValProGlyIleProThrAlaCysIleThrVal 236
Db 64 ACAGGTATTCAGGAATACAGAGATGCGCTGCCCAAGATCCAAACGCTGTATTACGGTGT 123
Qy 237 GlnAspAlaGluMetMetSerArgMetAlaSerHisGlyIleValIleGlnLeu 256
Db 124 GAAATATGCAAAATGATGTCAGAAATGGCTTCTCATGGGATCAAAATGTCATTCAGCTA 183
Qy 257 LysMetGlyAlaValThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThr 276
Db 184 AAGATGGGGGCAAGACCTACCCAGATCTGATTCCTTCAACACTGTAGCAGATCACT 243
Qy 277 GlySerIleTyrProGlnGlnValValLeuValSerGlyHisLeuAspSerTyrAspVal 296
Db 244 GGGAGCAAAATATCCAGAACAGTGTGTACTGGTCAGTGAATCTGACAGCTGGGATGTT 303
Qy 297 GlnGlnGlyAlaMetAspAspGlyGlyGlyAlaPheIleSerTyrGlnAlaLeuSerLeu 316
Db 304 GGGCAGGGGTGCATGATGATGCGGTGAGCGCTTATATCATGGGAGCACTCTCACCTT 363
Qy 317 IleLysAspLeuGlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThrAlaGlu 336
Db 364 ATTAAAGATCTTGGCTGCTGCTCCAAAGAGACTCTGCGGCTGGTCTCTGAGCTGCAAA 423
Qy 337 GlnGlnGlyGlyValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsn 356
Db 424 GAACAAGGTGGAGTGGTGGCTTCACGATATTATCATGACAAAGTAAATATTTCCAAAC 483
Qy 357 TyrSerLeuValMetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThr 376
Db 484 TACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
Qy 377 GlySerGlyLysAlaArgAlaIleMetGlnGlnValMetSerLeuLeuGlnProLeuAsn 396
Db 544 GGCAGTGAAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 603
Qy 397 IleThrGlnValLeuSerHisGlyGlyGlyIleThrAspIleAsnPheTyrIleGlnAlaGly 416

```

```

Db 604 ATCACTCAGGTCTGAGCCATGAGGAAGGACAGACATCAACTTTTGATCCAGCTGA 663
Qy 417 ValProGlyAlaSerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisSerHis 436
Db 664 GTGCTTGAGGCACTCTTACTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 723
Qy 437 Gly-AspThrMet 440
Db 724 GAAAGACACCATG 736

```

RESULT 13
LOCUS BX416896
DEFINITION BX416896 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE001YF04
5-PRIME, mRNA sequence.
ACCESSION BX416896
VERSION BX416896.1 GI:30650313
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE001DC02Q1&cluster=4663.f. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE001DC02Q1.

FEATURES
source
location/Qualifiers
1..1012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE001YF04"
/issue_type="PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 281 a 240 c 275 g 215 t 1 others
ORIGIN

Alignment Scores:

Pred. No.:	6.44e-134	Length:	1012
Score:	1252.00	Matches:	251
Percent Similarity:	99.21%	Conservative:	0
Best Local Similarity:	99.21%	Mismatches:	1
Query Match:	51.80%	Indels:	0
DB:	13	Gaps:	0

US-09-745-763-36 (1-472) x BX416896 (1-1012)

```

Qy 1 MetLysPheLeuIlePheAlaPheGlyGlyValHisLeuLeuSerLeuCysSerGly 20
Db 255 ATGAATTCCTTATCTTCGATTTTCGATGATGATGATGATGATGATGATGATGATGATGAT 314
Qy 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGlnGlnIleGlnIle 40
Db 315 AAAGCTATATGCAAGAAATGCGATCTTAAGAGGACTTTTGAAGAAATTAAGAAATA 374
Qy 41 Alasercysglaaspvalalalyalalileleasndualavaltyrglylyalagln 60

```



```

Db      375 GCCACGTGGAGATGTTCTTAAGCAATCATCACTGATTGTTATGTAAAGCCG 434
Qy      61 AenAserTyrGluArgLeuAlaLeuLeuValaAspThrValGlyProArgLeuSerGly 80
Db      435 AACCATCTTAAGACGATGGACCTTCTGTTGTAACCTGGAGACCCAGCTGAGTGGC 494
Qy      81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
Db      495 TCCAAAGAACCTTGAAGAAAGCCATCCAAATTAATGTAACCAAACTCAGCAAGATGGCTG 554
Qy      101 GlnLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGlnIleSerAla 120
Db      555 GAGAAAGTTCACTGAGAGCCAGCTGAGAAATACCCCACTGGAGAGGGAGAAATCAGCT 614
Qy      121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db      615 GGATGCTGGAGCCAGAAATTCATAGATGAGCCATCTGGGCTTGGCAGCAGCATTTGGG 674
Qy      141 ThrProGluGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160
Db      675 ACTCTTCAGAAAGCATTAACAGCAAGATTCGTGGTGTACCTCTTTCATTAAGTGCAG 734
Qy      161 ArgArgAlaSerGluAlaArgGlyLysIleValIleValTyrAsnGlnProTyrIleAsnTyr 180
Db      735 AGAAGGGCTCAGAAAGCAAGAGGAAAGATTGTTGTTATTAACCAACTTACATCACTAC 794
Qy      181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLe 200
Db      795 TCAGAGAGCGTGCAATACCGAAGCGAGGGGGCGGTGAGAGCTGCCAAGGTTGGGGCTTT 854
Qy      200 ValAspSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisTyrGlyIleG 220
Db      855 GGCACTCTCATTCATCCGTGGCTCTCTCTTCATCTACATCTCTCACACAGGATTTCA 914
Qy      220 nGluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaG 240
Db      915 GGAATATACAGAGATGGGTGCCCAAGATTCCAACAGCTGTATTACGGTGAAAGATGCA 974
Qy      240 UMetSerArgMetAlaSerHisGlyIleLysIle 252
Db      975 AATGATGTCAAGATGGCTTCTCATGGATCAAAATT 1011

RESULT 14
LOCUS   BX355939/c                      996 bp      mRNA       linear      EST 05-MAY-2003
DEFINITION   BX355939 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1006Y115 3-PRIME, mRNA sequence.
ACCESSION   BX355939
VERSION     BX355939.1 GI:30382027
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Buxaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE   1 (bases 1 to 996)
AUTHORS    Li W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished
COMMENT     Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1006AE08NPLcluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1006AE08NPL.
Location/Qualifiers
1..996
FEATURES
Source

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1006Y115"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      270 a      231 c      223 g      269 t      3 others
ORIGIN
Alignment Scores:
Pred. No.:      5,31e-133      Length:      996
Score:          1244.00      Matches:      241
Percent Similarity: 99.59%      Conservative: 0
Best Local Similarity: 99.59%      Mismatches: 1
Query Match:      51.47%      Indels:      0
DB:              13      Gaps:      0
US-09-745-763-36 (1-472) x BX355939 (1-996)
Qy      231 ThrAlaCysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGlyIle 250
Db      995 ACAGCTTATTTACGGTGGAAGATGCAAGAAATGATGTAAGAAATGCTTCATGGGATC 936
Qy      251 LysIleValIleGlnIleLysMetGlyValAlaLysThrTyrProAspThrAspSerPheAsn 270
Db      935 AAAATTGTCATTCAGCTTAAGATGGGGGCAAGACCTTAACCAATCTGATTCCTTCAAC 876
Qy      271 ThrValAlaGluIleThrGlySerLysTyrProGluGlnValLeuValSerGlyHis 290
Db      875 ACTGTAGCAGATCATCTGGAGCAAAATATCCAAACAGTTGATCTGCTGACGGACAT 816
Qy      291 LeuAspSerTyrAspValGlyGlnGlyAlaMetAspAspGlyGlyAlaPheIleSer 310
Db      815 CTGAGACGCTGGATGTTGGGCAAGGTCATGATGATGGCGGTGAGCCTTATATATCA 756
Qy      311 TrrGluAlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeuArgLeu 330
Db      755 TGGAGACACTCTCACTTATTAAGATCTGGGCTGGCTCCAAAGAGACTTCGGCTG 696
Qy      331 ValLeuTrrThrAlaGluGluGlnGlyValGlyAlaPheGlnTyrTyrGlnLeuHis 350
Db      695 GTGCTCTGAGCTGCAGAAAGAACAGTGAGCTGTGCTCTTCCAGTATTATCACTTACAC 636
Qy      351 LysValAsnIleSerAsnTyrSerLeuValMetGluSerAspAlaGlyThrPheLeuPro 370
Db      635 AAGGTAAATATTTCCACTACAGCTGTGATGAGTCGACGACGAACTTCTTACC 576
Qy      371 ThrGlyLeuGlnPheThrGlySerGlyLysAlaArgAlaIleMetGluGluValMetSer 390
Db      575 ACTGGGTGCATATTCATCTGGCAGATGAAGGCCAGGCCATCATGAGAGAGTTATAGC 516
Qy      391 LeuLeuGlnProLeuAsnIleThrGlnValLeuSerHisGlyGlnGlyThrAspIleAsn 410
Db      515 CTGCTGAGCCCTCAATATCTCAAGTCTTGAGCATGAGAGGAGACAGACATCAAC 456
Qy      411 PheTrrIleGlnAlaGlyValProGlyAlaSerLeuAspAspLeuTyrLysTyrPhe 430
Db      455 TTTTGGATCCAAAGTGGAGTGCTCGAGCCAGCTTACTTATGACTTATACAAATATTC 396
Qy      431 PhePheHisHisSerHisGlyAspThrMetThrValMetAspProLysGlnMetAsnVal 450
Db      395 TTTCTTCATCACTCCACGAGACACATGATCTGTCTCATGTAGTCCAAAGCAGATAAAGTT 336
Qy      451 AlAlaAlaValITrrPalValIleValSerTyrValValAlaAspMetGluGluMetLeuPro 470
Db      335 GCTCTGCTGTTTGAGGCTGTTGTTCTTATGTTGTTCAGACATGAGAAATGCTGCT 276
Qy      471 ArgSer 472
|||||

```

Db 275 AGTCC 270

RESULT 15
BO938234

LOCUS BO938234 926 bp mRNA linear EST 21-AUG-2002

DEFINITION AGENCOURT 8932063 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6465955

ACCESSION BO938234

VERSION BO938234.1 GI:2253712

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL).
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM13989 row: h column: 20
High quality sequence start: 38
High quality sequence stop: 605.
Location/Qualifiers
1..926
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6465955"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

BASE COUNT 252 a 200 c 259 g 215 t

ORIGIN

Alignment Scores:

Pred. No.:	1,86e-129	Length:	926
Score:	1213.00	Matches:	254
Percent Similarity:	89.93%	Conservative:	14
Best Local Similarity:	85.23%	Mismatches:	24
Query Match:	50.19%	Indels:	6
DB:	13	Gaps:	0

US-09-745-763-36 (1-472) x BO938234 (1-926)

Qy 37 LyeGluGluIleAlaSerCysGlyAspValAlaIleValIleIleAsnLeuAlaValTyr 56

Db 22 CGTGGGAAATAGCCAAATGATGATGCTTAAGCAATTATCACTGCTGTAT 81

Qy 57 GilyysAlaGlnAsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyPro 76

Db 82 GGTAAATACCAAGACCGGCTTATAGCGTTGGACCTTGTAGTACTGTGGACCC 141

Qy 77 ArgLeuSerGlySerLeuAsnLeuGluValAlaIleGlnIleMetTyrGlnAsnLeuGln 96

Db 142 AGACTGATGGCTTAAAGAACTTAGAAAGCTATCAATCATATACCAAACTGCAAG 201

Qy 97 GlnaArgIleuGluValAlaIleuGluProValArgIleProHisTyrGluArgGly 116

Db 202 CAAGATGGGCTTGAAGAAATGTTCACTGGAGCAGTCAAGATACCCCACTGGGAGAGGGA 261

Qy 117 GluGluSerAlaValMetLeuGluProArgIleHisIleValIleIleLeuGlyLeuGly 136

Db 262 GAAGATCTGAGAGAGCTTGAAGCCCGAATTACAAAGATGCTATTACAGCTTGGC 321

Qy 137 SerSerIleGlyThrProProGluGlyIleThrAlaGluValLeuValIleThrSerPhe 156

Db 322 AGCAGCATTTGGGACTCCCGCAGAGGAGCATCAAGCAGAAAGTGGTGGTAGCCCTTTT 381

Qy 157 AspGluLeuGlnAAGAGAlaSerGluAlaArgGlyIleValIleValIleTyrAsnGlnPro 176

Db 382 GATGAACCTTCAAGAAAGAGCATTCAGAGCAAGAGGAGATCATTTTATACCACT 441

Qy 177 TyrIleAsnTyrSerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaIleLys 196

Db 442 TACACTGGCTATGAGAAAGCTGTGCATACCGGCTGAGGAGCTGTGAAGCTGCCAG 501

Qy 197 ValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHis 216

Db 502 GTGGGAGCTGTGGCATCCATCATCAATCAATGATAGCTTCTTTCATCTACAGTCTCAC 561

Qy 217 ThrGlyIleGlnGluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrVal 236

Db 562 ACAGGTATTCAGAAATATCAAGATGATGTGCCCAAGATTCACAGCTGTATCACAGTA 621

Qy 237 GluAspAlaGluMetMetSerArgMetAlaSerHisGlyIleValIleGlnLeu 256

Db 622 GAATGTCAGAAATGATGTCTCCAAATGGCTTCTGTGGGAGCAAAATGTCATTCACTG 681

Qy 257 LysMetGlyAlaIleTyrThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThr 276

Db 682 GAAATGGAGCAAGAACCTTATCCAGATATGATTCCTTCATATCTGTTCAGAGATCACT 741

Qy 277 GlySerIleTyrProGluGlnValVal-LeuValSerGlyHisLeuAspSerTyrAsp 296

Db 742 GGGAGCATGTATCCAGAGAAAGTTGCTCGTTCAGTGAACATTGACACGCTGGGAGTG 801

Qy 296 AlGlyGlnGlyAlaMetAsp-AspGlyGlyGlyAla-PheIleSerTyrGlu-AlaLeu 315

Db 802 TTGGGCAAGGTGCACTGGAATGAGGGGGGGAACCTTTCATATCATGGGAAGCACTCT 861

Qy 315 er-LeuIleLysAspLeuGlyLeuArgProLysArgThrLeu 328

Db 862 CCTTGTAAAGATCTTGGGGCTGGCCCCCAAGAGAGACTC 903

Search completed: December 23, 2003, 01:44:34
Job time : 2358 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 22, 2003, 23:17:31 ; Search time 95 Seconds
(without alignments)
2192.977 Million cell updates/sec

Title: US-09-745-763-36
Perfect score: 2417
Sequence: 1 MKFLIPAFGCVHLLSLCSG.....AVMAVSVYVADWEMLPRS 472

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=x1p
-Q=/cgn2_1/USPTO.spool_p/US09745763/runat_22122003_113318_25275/app_query.fasta_1.647
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=ptc -NOEM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09745763 -OCGN_1_1_85 -runat_22122003_113318_25275 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2417	100.0	1863	4	US-09-482-273-28
2	1338	55.4	1134	4	US-09-482-273-95
3	229.5	9.5	1491	4	US-09-079-955-1
4	225.5	9.3	1640	3	US-09-330-095-2
5	182	7.5	1644	4	US-09-252-991A-2005
6	182	7.5	1755	4	US-09-252-991A-1721
7	166	6.9	2653	1	US-08-325-553-1
8	166	6.9	2653	2	US-08-394-152A-1
9	166	6.9	2653	4	US-08-705-477E-1
10	161	6.7	2133	4	US-09-164-034B-1
11	161	6.7	2387	4	US-08-705-477E-100
12	149	6.2	765	4	US-09-252-991A-1872

13	149	6.2	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
14	149	6.2	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
15	127	5.3	4080	4	US-09-016-434-1352	Sequence 1352, Ap
16	127	5.3	5010	3	US-09-161-244-1	Sequence 1, Appl1
17	126	5.2	780	1	US-08-325-553-27	Sequence 27, Appl
18	126	5.2	780	2	US-08-394-152A-27	Sequence 27, Appl
19	126	5.2	780	4	US-08-705-477E-27	Sequence 27, Appl
20	108.5	4.5	11935	4	US-09-634-238-401	Sequence 401, App
21	105	4.3	1230025	4	US-09-198-452A-1	Sequence 1, Appl1
22	103	4.3	7470	3	US-08-417-089-5	Sequence 5, Appl1
23	103	4.3	7470	3	US-08-695-651-5	Sequence 5, Appl1
24	103	4.3	7470	3	US-08-330-285-5	Sequence 5, Appl1
25	103	4.3	7470	3	US-08-695-421-5	Sequence 5, Appl1
26	103	4.3	7470	4	US-08-697-826A-9	Sequence 9, Appl1
27	102.5	4.2	1014	4	US-09-107-532A-2344	Sequence 2344, Ap
28	100.5	4.2	1419	1	US-08-242-098-39	Sequence 39, Appl
29	100.5	4.2	1974	4	US-09-423-439-59	Sequence 59, Appl
30	100.5	4.2	3177	4	US-09-252-991A-521	Sequence 521, App
31	100.5	4.2	80161	3	US-09-036-987A-1	Sequence 1, Appl1
32	100.5	4.2	80161	3	US-09-370-700-1	Sequence 1, Appl1
33	100.5	4.2	80161	4	US-09-603-207-1	Sequence 1, Appl1
34	100	4.1	36181	4	US-08-311-731A-120	Sequence 120, App
35	98.5	4.1	2805	4	US-09-252-991A-10208	Sequence 10208, A
36	98.5	4.1	4278	4	US-08-661-527-153	Sequence 153, App
37	98.5	4.1	124884	4	US-09-661-596A-76	Sequence 76, Appl
38	98	4.1	1847	4	US-09-321-017B-998	Sequence 998, App
39	97.5	4.0	1368	4	US-09-328-352-1713	Sequence 1713, Ap
40	97.5	4.0	1860	4	US-09-107-532A-1243	Sequence 1243, Ap
41	97.5	4.0	2048	3	US-08-776-246-1	Sequence 1, Appl1
42	97.5	4.0	2048	3	US-08-776-251-1	Sequence 1, Appl1
43	97	4.0	3300	1	US-08-194-290-6	Sequence 6, Appl1
44	97	4.0	3300	2	US-08-614-377A-6	Sequence 6, Appl1
45	97	4.0	3300	3	US-09-142-648B-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-482-273-28
; Sequence 28, Application US/09482273
; Patent No. 6534631
; GENERAL INVENTION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; EARLIER FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 1863
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-482-273-28
Alignment Scores:
Pred. No.: 8.56e-277
Score: 2417.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 4
Gaps: 0
US-09-745-763-36 (1-472) x US-09-482-273-28 (1-1863)
Length: 1863
Matches: 472
Conservative: 0
Mismatch: 0
Indels: 0

```

Qy 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20
Db 99 ATGAAATTCCTTATCTTCGATTTTTCGGTGTGTTCACCTTTATCCCTGTGCTGG 158
Qy 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluIleLysGluIle 40
Db 159 AAACATATATCAAGAAATGGCATCTCTTAAGAGCATTTTGAAGAAATTAAGAAATA 218
Qy 41 AlaserCySerGlyAspValAlaLysAlaIleLeuLeuAlaValIleGlyValAlaGln 60
Db 219 GCCAGCTGTGAGAGTGTGTAAAGCAATCAATCACTAGCTGTTTATGTAAAGCCAG 278
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db 279 AACGATCTCATAGAGCATTTGGCACTTCTGGTGTATCTGTGGACCCAGACTGAGTGGC 338
Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100
Db 339 TCCAAGAACTTGAAGAAAGCCATCCAAATTTGTACCAAAACCTGCACAGATGGGCTG 398
Qy 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerAla 120
Db 399 GAGAAAGTTCACTGTGAGCCAGTGAATACCCCACTGGAGAGGAGGAAAGAAATCAGCT 458
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 459 GTGATGCTGGAGCCAGAAATTCATAGATAGCAATCCCTGGGTCTTGGACAGCATGGG 518
Qy 141 ThrProProGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160
Db 519 ACTCTCCAGAAAGGATTAACAGAGAACTTCTGGGTGTGAGCACTTTTGATGAATCAG 578
Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValIleTyrAsnGlnProTyrIleAspTyr 180
Db 579 AGAAGGGCTTCAAGAAAGAGAGAGAGTGTGTAAACCAACTTACATCAACTAC 638
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200
Db 639 TCAAGAGAGGTGCATACCAAGACGAGGGGCGGTGAGAGCTGCCAAGTTGGGCTTG 698
Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db 699 GCATCTCTCATTCGATCCGTGGCTCTTCTCACTCAAGCTCCTCAACAGATTCAG 758
Qy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGln 240
Db 759 GAATACCGAGATGGGTGCCCAAGATTCCAACAGCTGTATTAACGTTGAAAGATGCAGAA 818
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 819 ATGATGTCAAGATATGCTTCTCATGTGGATCAAAAATGTTCATTCAGCTAAAGATGGGCA 878
Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
Db 879 AAGACTTACCAAGATGATGATCTCTCAACACTGTAGACAGATCATCTGGAGCAAAATAT 938
Qy 281 ProGluGlnValIleValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300
Db 939 CCAAGAACGGTTGATCTGTGATGACATCTGCACACTGGAGATGGTGGAGGGGTGCC 998
Qy 301 MetAspAspGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
Db 999 ATGATGATGCGCGGTGAGCTTTATATCATGTGGAGACATCTCATTTTAAAGATCTT 1058
Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThrAlaGluGluGlnGly 340
Db 1059 GGGGCTGTCCAAAGAGACTCTGGGCTGGGTCTGCACTGCAGAGAAACAAGTGGGA 1118
Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAlaIleSerAsnTyrSerLeuVal 360
Db 1119 GTTGATGCTTCCAGATATTAATCAAGTAAACAAGGTAATATTTCCAACTTACAGTCTGTG 1178
Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380

```

```

Db 1179 ATGAGTCTGACGAGAACCTTCTTACCCATCGGCTGCATATTCATCGCAGTGAAG 1238
Qy 381 ALAArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAlaIleThrGlnVal 400
Db 1239 GCCAGGGCCATCATGAGAGAGGATTAAGCTGTGTGAGCCCTCCAAATATCACTCAGTGC 1298
Qy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420
Db 1299 CTGAGCCATGAGAGAGGACACATCATCTTTTGATCTCAAGCTGAGAGTCCGTGAGCC 1358
Qy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
Db 1359 AGTCTACTGATGACTTATACAGATATTTCTTCCATCATCTCCACGAGACACCATG 1418
Qy 441 ThrValMetAspProLysGlnMetAspValAlaAlaAlaValTyrPheAlaValSerTyr 460
Db 1419 ACTGTCAATGATCCAAAGACAGATGAATGTGTGCTGTGTGGCTGTGTGTTTCTTAT 1478
Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db 1479 GTTGTTCAGATGAGAGAAATGCTGCTTAGTTC 1514

RESULT 2
US-09-482-273-95
Sequence 95, Application US/09482273
Patent No. 6534631
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
EARLIER FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1998-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 95
LENGTH: 1134
TYPE: DNA
ORGANISM: Homo sapiens
US-09-482-273-95

Alignment Scores:
Pred. No.: 5,07e-149 Length: 1134
Score: 1338.00 Matches: 261
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 55.36% Indels: 1
DB: 4 Gaps: 0

US-09-745-763-36 (1-472) x US-09-482-273-95 (1-1134)
Qy 211 SerIleTyrSerProHisThrGlyIleGlnGluTyrGlnAspGlyValProLysIlePro 230
Db 1 TCCATCTACAGTCTCTCAACAGGATTCAGGAATACCAAGATGGCGGCCAAGATTCGA 60
Qy 231 ThrAlaCysIleThrValGluAspAlaGluMetCysArgMetAlaSerHisGlyIle 250
Db 61 ACAAGCTGTATTTAGCGTGAAGATGACGAATGATGTCAAGAAATGAGGCTTCTCATGGGATC 120
Qy 251 LysIleValIleGlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPheAsn 270
Db 121 AAAATTTCAATCACTGTAAGATGGGGCAAGACCTTACCAAGATGATGATCTTCAAC 180
Qy 271 ThrValAlaGluIleThrGlySerLysTyrProGluGlnValLeuValSerGlyHis 290

```

```

Db      181 ACTGACAGAGATCACTGGAGCAATATCCAGAACAGTTGTAAGTGTGATGAGCAT 240
Qy      291 LeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGlyGlyGlyAlaPheIleSer 310
Db      241 CTGGACAGCTGGAGTGTGGGCGAGGTGTCATGATGATGAGGAGGAGCTTTATATCA 300
Qy      311 TTPGluAlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgTrpLeuArgLeu 330
Db      301 TGGAGACACTCTCATTTATTAAGATCTTGCGCTGCTGCTCAAGAGAGAGACTGCGGCTG 360
Qy      331 ValLeuTrpThrAlaGluGlnGlnGlyGlyAlaGlyAlaPheGlnTrpTrpGlnLeuHis 350
Db      361 GTGGCTGCACTGCAGAGAGAACAGAGTGGAGTTGGCTTCCAGTATTTACAGTTACAC 420
Qy      351 LysValAsnIleSerAsnTrpSerLeuValMetGluSerAspAlaGlyTrpPheLeuPro 370
Db      421 AAGGTAATATTTTCAACTACAGTCTGTGATGAGTGTGACGAGAGAACCTTTTACC 480
Qy      371 ThrGlyLeuGlnPheThrGlySerGlyLysValAlaArgAlaIleMetGluGluValMetSer 390
Db      481 ACTGGGCTGCATTCATCTGGCAGTGAAGAGCCAGGCG-ATCATGAGAGAGGTTATGAC 539
Qy      391 LeuLeuGlnProLeuAsnIleThrGlyValLeuSerHisGlyGlnGlyTrpAspIleAsn 410
Db      540 CTGGCTGAGCCCTCAATATCATCTCAGGTCCTGAGCATGAGAGAGGACAGACATCAAC 599
Qy      411 PheTrpIleGlnAlaGlyValProGlyAlaSerLeuAspAspLeuTrpLysTrpPhe 430
Db      600 TTTTGGATCCAAAGTGAAGTGCCTGGAGCCAGTCTACTGATGACTTATTAACAATTTTC 659
Qy      431 PhePheHisHisSerHisGlyAspThrMetThrValMetAspProLysGlnMetAsnVal 450
Db      660 TTCTTCATCATCTCCACGAGAGACACATGCTGCTCATGATCCAAAGCAGATGAAGTT 719
Qy      451 AlaAlaAlaValTrpAlaValAlaSerTrpValValAlaAspMetGluGluMetLeuPro 470
Db      720 GCTGCTGCTGTTGGCGTGTGTTCTTATGTTGTTCGACATGAGAAATGCTGCT 779
Qy      471 ArgSer 472
Db      780 AGGTCC 785

```

RESULT 3

```

US-09-079-955-1
; Sequence 1, Application US/09079955A
; Patent No. 6465209
; GENERAL INFORMATION:
; APPLICANT: Alexander Blinkovsky
; APPLICANT: Kimberly Brown
; APPLICANT: Elizabeth Gollighly
; APPLICANT: Tony Byun
; APPLICANT: Thomas Mathiasen
; APPLICANT: Lene V. Kolof
; APPLICANT: Mikio Fujii
; APPLICANT: Chigusa Shizuka
; TITLE OF INVENTION: Methods For Producing Protein
; TITLE OF INVENTION: Hydrolyses
; FILE REFERENCE: 5253.500-US
; CURRENT APPLICATION NUMBER: US/09/079.955A
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
US-09-079-955-1

```

```

Alignment Scores:
Pred. No.: 3.39e-17 Length: 1491
Score: 229.50 Matches: 96
Percent Similarity: 42.12% Conservative: 83
Best Local Similarity: 22.59% Mismatches: 175

```

```

Query Match: 9.50% Indels: 71
DB: 4 Gaps: 19
US-09-745-763-36 (1-472) x US-09-079-955-1 (1-1491)
Qy      77 ArgLeuSerGlySerLysAsnLeuGluValAlaIleGlnIleMetTrpGlnAsnLeuGln 96
Db      157 CGCGTCTTGTGTGTAAGCCACAGCAGACAGCTTAATCATCTTACGAGAGCTGAAG 216
Qy      97 GluAspGlyLeuGluLysValHisLeuGluProValArgIleProHisTrpGluArgGly 116
Db      217 AAGACTGGCTACTATGATGATGTTACAGACAGCT--CAGGTGACCTGTGAGCAATGCC 273
Qy      117 GluGluSerAlaValMetLeuGluProArgIleHisIleValAlaIleLeuGlyLeuGly 136
Db      274 GACCAGACGCTCAAGTGGCGGCGATGAGGAATGAC-----GCCAAG 315
Qy      137 SerSerIleGlyThrProProGluGlyIleThrAlaGluValLeuValAlaTrpSerPhe 156
Db      316 ACCATGACCTACAGTCCAGCGTCAGGTCACCGCGATGACCGCTTCAAGAACCTG 375
Qy      157 AspGluLeuGlnArg--ArgAlaSerGluAlaArgGlyLysIleValAlaTrpAsnGln 175
Db      376 GGAATGACGAGGCGGATTAACCATCGATGCGAGGCGAAGGTGCGCTGATCAACGCT 435
Qy      176 -----ProTrpIleAsnTrpSerArgTrpValGlnTrpArgTrpGlnGlyAlaVal 192
Db      436 GGAATGCGCCGTTCCGCGCAAGTCC-----GTT 465
Qy      193 GluAlaAlaLysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIle 212
Db      466 CTGGCTGCCAAGGCAAGCGCGCGGCTTGATTTGATTAACAATGCGCGGATTCATG 525
Qy      213 TyrSerProHisThrGlyIleGlnGlyTrpGlnAspGlyValProLysIleProThrAla 232
Db      526 GCGGACACCTTGAGCGCGGCGAG-----AGTATAGAGACCGTATTCGCCATTGTC 579
Qy      233 CysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGlyIleLysIle 252
Db      580 GGTATCAGCTTGAGAGATGCGCAGAACCTGATCAAGCTTCTGAGGTGAG--TCGGA 636
Qy      253 ValIleGlnLeuLysMetGlyAlaLysTrpTrpProAspThrAspSerPheAsnTrpVal 272
Db      637 TCGTGTGATCTGCGGATGATGATGAAGCAGAGAACCTGAG--ACGTATTAAGTTGTC 693
Qy      273 AlaGluIleThrGlySerLysTrpProGluGlnValAlaLeuValSerGlyHisLeuAsp 292
Db      694 GCGCAGACGAAAGGCGCGCAT--CCGAAACAAGTCTGCGCGTGGCGCACACGAG 750
Qy      293 SerTrpAspValGlyGlnGlyAlaMetAspAspGlyGlyGlyAlaPheIleSerTrpGlu 312
Db      751 TCAGTCAGCGCGGCGCTGTGATACACAGATGAGCTCGGCGCATTTATGACACTTGGTC 810
Qy      313 AlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgTrpLeuArgLeuValLeu 332
Db      811 ATTGCCAAGCGCTCAGCAGACTCTCGTC-----AAGATGCCGCGCTCTCTTC 864
Qy      333 TrpThrAlaGluGlnGlnGlyGlyAlaGlyAlaPheGlnTrpTrpGlnLeuHis 347
Db      865 TGGACACAGAGAGGAGTTCGCTGCTGCGGAGC--AACTACTACGTTCCATCTGAAT 921
Qy      348 -----GlnLeuHisLysValAsnIle-----SerAsn 356
Db      922 GCCACGAGCTGAACCAAGATCCAGCTTACTGAACTTGCACATGATGCGCTCACTTAAC 981
Qy      357 TyrSerLeu--ValMetGluSerAspAlaGlyTrpPheLeuProThrGlyLeuGlnPhe 375
Db      982 TAGCGCTCATGATCTATGAAGATGATGATGCGCGCTTCAACAGAGAGGAGCGC----- 1035
Qy      376 ThrGlySerGlyLysValAlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeu 395
Db      1036 GCCGTTTCGCCGCCAGATCGAGAAACTGTCGAGAC-----TACTACGACTCCATC 1086

```



```

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2005
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2005

Alignment Scores:
Pred. No.: 1,74e-11      Length: 1644
Score: 182.00           Matches: 95
Percent Similarity: 38.71%      Conservative: 61
Best Local Similarity: 23.57%    Mismatches: 141
Query Match: 7.53%             Indels: 106
Gaps: 19

US-09-745-763-36 (1-472) x US-09-252-991A-2005 (1-1644)

QY 65 GluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGlySerIysAsnLeu 84
Db 1395 GAGAGACATCGCACTCTCAAGAC-----GGCAACCGCGCGCGCGACCGCGGGCTAC 1342
QY 85 GluLysAlaIleGlnIleMetTYrGlnAsnLeuGlnIleAspGlyLeuGluValHis 104
Db 1341 CAGGCTTCGCTCGACTAGTAAAGCAGCCTCGCAAGAACCGGCTAC---AAGTCAAGC 1285
QY 105 LeuGluProValaLargIleProHisIleProLargIleGluGluSerAlaValMetLeuGlu 124
Db 1284 GTGAGCGCTTCCTCCGCTACCGCTAC-----TAC 1255
QY 125 ProArgIleHisIysIleAlaIleLeuGlyLeuSerSerIleGlyThrProGlu 144
Db 1254 CCGAAA-----GGCCGGGTACCTGAGCGGACCGCTCCGCG 1216
QY 145 -----Gly 145
Db 1215 CCGGTACCTAGAGATGAGAGAAAGACTTCACTGTCGACAGCGAGCGAGCGAC 1156
QY 146 IleThrAlaGluValIleValVal-----ThSer 155
Db 1155 GTCAACCGCAAGGTGTCCTCGGTGACCTGTCTCGCGCGCGCAACACTCCACAGC 1096
QY 156 PheAspGluLeuGlnArgAlaSerGluAlaArgGlyIysIleValIleValIleAsnGln 175
Db 1095 GGTTCGAGAGGAGAACTTCCCACTTCCCGCGCGCTGATCGCGTATATCCAGCGC 1036
QY 176 ProTYrIleAsnTYrSerArgThrValGlnTYrArgThrGlnGlyAlaValGluAla 195
Db 1035 GGCACCTGCACCTTCAG-----CAGAAAGCGGAGAACGCCGCG 997
QY 196 LysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTYrSerPro 215
Db 996 GCCCGCGCGCGCGCGGTGATCATCTTCAACAGCGGCAAC-----ACCGAGCAGC 946
QY 216 HisThrGlyIle-----GlnGluTYrGlnAspGlyValProIle 229
Db 945 CGCAAGGCGCTCGAAGACGTACCGTGGCGGAGTCCTTCAGAGGCGCGATCCCGGTATC 886
QY 230 ProThrAlaCysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGly 249
Db 885 -----TTGCGCACCTAGCAACAGCGCTGCGCTGTGCGACAGCCCGAGCTCGAG 835
QY 250 IleLysIleValIleGlnIleLeuMetClyAlaIleTYrProAspThrAspSerPhe 269
Db 834 TTGCACTGTGTGTGACGTG-----GTACGCAAGAAAGACCGAGACCTTAC 790

```

```

QY 270 AsnThrValAlaGluIleThrGlySerIysTYrProGluGlnValIleValSerGly 289
Db 789 AACGTGTCCCGGAG---ACCGCGCGGCAACCCGAACAGTGTGATGTGCGCGG 733
QY 290 HisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGlyGlyAlaPheIle 309
Db 732 CACTTCACCTCGGTGTGGAAGGCCCGGTATCAACAGATTCGGGCGAGCGCGC 673
QY 310 SerTrpGlu---AlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeu 328
Db 672 CACTGAGATATGCGGTGCTGCTGCGCAAGCGCTG-----CCGTCAACAGAGTG 622
QY 329 ArgLeuValLeuTrpThrAlaGluGlnGlyGlyValGlyAlaPheGlnTYrGln 348
Db 621 CGTTGCGCTGTGTGGCGCGCGAGAACCGCGCTGTGGCTGAGCCACTATGTCAG 562
QY 349 LeuHisLysValAsnIleSerAsnTYrSerLeuValMetGluSerAspAlaGlyThrPhe 366
Db 561 -----ACCTCGCGCGGAGAG----- 544
QY 369 LeuProThrGlyLeuGlnPheThrGlySerGluLysAlaArgAlaIleMetGluGluVal 388
Db 543 -----AAGAAAGATCAAGGCTTACCTGAAC----- 517
QY 389 MetSerLeuLeuGlnProLeuAsnIleThrGlnValLeuSerHisGlyGlnGlyThrAsp 408
Db 516 TTGACATATGTCGCTCGCGCACTTGCACATTATATAGACGGAGCGGTTCGAC 457
QY 409 IleAsnPheTrpIleGlnAlaGlyValProGlyAlaSerLeuLeuAspAspLeuTYrLys 428
Db 456 -----TTCCGCTCCGAG---GGTCCCGCGCTCGCGCCCATCGAGCCCTGTTGAA 406
QY 429 TyrPhePhe 431
Db 405 GCCTACTTC 397

RESULT 6
US-09-252-991A-1721
; Sequence 1721, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1721
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1721

Alignment Scores:
Pred. No.: 1.93e-11      Length: 1755
Score: 182.00           Matches: 95
Percent Similarity: 38.71%      Conservative: 61
Best Local Similarity: 23.57%    Mismatches: 141
Query Match: 7.53%             Indels: 106
Gaps: 19

US-09-745-763-36 (1-472) x US-09-252-991A-1721 (1-1755)

QY 65 GluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGlySerIysAsnLeu 84
Db 367 GAGGACATCGCACTCTCAAGAC-----GGCAACCGCGCGCGCGACCGCGGGCTTAC 420
QY 85 GluLysAlaIleGlnIleMetTYrGlnAsnLeuGlnIleAspGlyLeuGluValHis 104

```

```

Db      421  CAGGCTCTCGCTAGTACGTGAAGCAGCCCTGCAAGAAAGCCGCTAC---AAGGTACGC 477
      105  LeuGluProValAlaGlyLeuProHisTrpGluArgGlyGluGluSerAlaValMetLeuGlu 124
      478  GTGAGCCCTTCCCTTCACCGCTAC-----TAC 507
      125  ProArgIleHisIleValIleLeuGlyLeuGlySerSerIleGlyThrProProGlu 144
      508  CCGAAA-----GACCGGGTACCTGAGCGCCACCTGCGCGAG 546
      145  -----Gly 145
      547  CCGGTACCTAAGATGGAGAGAGACTTACACTGCTGCGAGACCGAGGAGGAGC 606
      146  IleThrAlaGluValLeuValAla-----ThSer 155
      607  GTACCCGCAAGGTGTCCCGGTGACCTGTCCGCGCGGAGACACCTCCACGAGC 666
      156  PheAspGluLeuGlnArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGln 175
      667  GGTTCGAGGGGAGAGACTTCCCACTTCCGCGCGGTGATGCGCTGATCCAGGCGC 726
      176  ProTyrIleAsnTyrSerArgThrValGlnTyrArgThrGlnGlyAlaValAlaAla 195
      727  GGCACTGCAACTTCGAG-----CAGAAGCGCGAGAGCCGCG 765
      196  LysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerPro 215
      766  GCCCGCGCGCGCGCGGTGATCATCTTCAACCGAGGCAC-----ACCGAGCAGC 816
      216  HisThrGlyIle-----GlnGluTyrGlnAspGlyValProLysIle 229
      817  CGCAAGGCGCTGGAACGTACCTGCGGCGAGCTCTACAGAGCGGCGGTGATC 876
      230  ProThrAlaCysIleThrValGluAspAlaGluMetMetSerArgMetAlaSerHisGly 249
      877  -----TTGCGCACTGACAGACAGCGCGGTGCTGTGCAACCCCGGACTGAG 927
      250  IleLysIleValIleGlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPhe 269
      928  TTGCACTGTGTGTGACGTG-----GTACGCAAGAAACCGAGACTTAC 972
      270  AsnThrValAlaGluIleThrGlySerLysTyrProGluGlnValValLeuValSerGly 289
      973  AACGTGTGCGGAG---ACCGTGTGCGGCAACCGAAACAACGTGTATGTGTGCGCGC 1029
      290  HisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGlyGlyAlaPheIle 309
      1030  CACCTCGACTGTGTGTGGAAGGCCCGGTATCAACGACACGTTGCGGCGAGCGCGC 1089
      310  SerTrpGlu---AlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeu 328
      1090  CAACGTGAGATGCGCTGTGCTGTGCGCAAGGCGCTG-----CCGTCACAAAGGTG 1140
      329  ArgLeuValLeuTrpThrAlaGluGlnGlnGlyValGlyAlaPheGlnTyrTrpGln 348
      1141  CGCTTGGCTGTGTGGGCGCGAGAGCGGCTGTGTGTGCAACCACTACCTGCGAG 1200
      349  LeuHisLysValAsnIleSerAsnTyrSerLeuValMetGluSerAspAlaGlyThrPhe 368
      1201  -----AACTCGCGCCGGAAGAG----- 1218
      369  LeuProThrGlyLeuGlnPheThrGlySerGluLysAlaArgAlaIleMetGluGluVal 388
      1219  -----AAGAAGAAATCAAGGCTTACTCTACAC----- 1245
      389  MetSerLeuLeuGlnProLeuAsnIleThrGlnValLeuSerHisGlyGlnGlyThrAsp 408
      1246  TTCGACATGATCGCTCGCCGCAACTTCGCAACTTATCATATGACGGCGGCTTCCGAC 1305
      409  IleAsnPheTrpIleGlnAlaGlyValProGlyAlaSerLeuLeuAspLeuTyrLys 428

```

```

Db      1306  -----TTGCGCTTCAG---GTTCCGCGCGGCTGCGCCGCTCATGAGCGCTGTTCGAA 1356
      429  TyrPhePhe 431
      1357  GCTTACTTC 1365

RESULT 7
US-08-325-553-1
; Sequence 1, Application US/08325553
; Patent No. 5538866
; GENERAL INFORMATION:
; APPLICANT: Israel, Ron S.
; APPLICANT: Heston, Warren D.W.
; APPLICANT: Fair, William R.
; TITLE OF INVENTION: THE PROSTATE-SPECIFIC MEMBRANE ANTIGEN
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,553
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/973,337A
; FILING DATE: 05 NOV 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 1747/41426
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)-977-9550
; TELEFAX: (212)-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Carcinoma
; IMMEDIATE SOURCE:
; CLONE: Prostate-Specific Membrane Antigen
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 262..2511
; US-08-325-553-1

Alignment Scores:
Pred. No.: 3,03e-09 Length: 2653
Score: 166.00 Matches: 105
Percent Similarity: 32.47% Conservative: 70
Best Local Similarity: 19.48% Mismatches: 169
Query Match: 6.67% Indels: 195
DB: 1 Gaps: 23

US-09-745-763-36 (1-472) x US-08-325-553-1 (1-2653)
      3 PheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuLeuCysSerGlyLysAla 22

```

```

Db      355  TTCTTTCTCTCGGCTCTCTCTGCGGTGTTTAAATCTCCAAATGAGCTACTAC 414
Qy      23    11eCysLysAngIYleSerIYleArgThrPheGluGluIleYsGluGluIleAlaSer 42
Db      415  ATTACTCCAAAGCAATATATGAAAGCATTTTGGATGATGAAAGCTGAG----- 465
Qy      43    CysGlyAspValAlaIleIleAsnLeuAlaValYrGlyLysAlaGlnAsnArg 62
Db      466  -----AACATCAAGAAAGTTCTTATATATTT----- 492
Qy      63    SerYrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGlySerLys 82
Db      493  -----ACACAGATACCAACATTTTACAGAGAAACGAA 522
Qy      83    -----AsnLeuGluLysAlaIleGlnIleMetYrGlnAsnLeuGlnAsnArg 99
Db      523  CAAAACTTTCAGCTTGCAAGCAAAATTCATCCAGTGAAAGAAATTTGGCTGGATTTCT 582
Qy      100   LeuGluLysValHisLeuGluProVal-----ArgIleProHisTrp 113
Db      583  GTTAGAGCTAGACACTTATGATGCTGTTGCTCTACCAATAAGACTCATCCCACTAC 642
Qy      114  -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
Db      643  ATCTCAATTAATTAAGAAATGAGAAATGATTTTCAACACATCATTTATTTGAACCA--- 699
Qy      127   IleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146
Db      700  -----CTCTCTCCAGGATAT 714
Qy      147   ThrIaGluValLeuValIleThrSerPheAspGluLeuGlnArgAlaSerGluAla 166
Db      715  GAAATGTTTCGATATGTAACACCTTCAGTCTTCTCTCTCAAGAAATG---CCA 771
Qy      167   ArgGlyLysIleValIleValYrAsnGlnProYrIleAsnYrSerArgThr----- 183
Db      772  GAGGCGCATCTAGTG-----TATGTTAACTATGACAGAACTGAAAGACTTC 816
Qy      184  -----ValGlnYrArgThrGlnGlyAlaValGluAlaIleVal 197
Db      817  TTTAAATTTGAAAGGAGCATGAAATCAATTCCTGGAAATTTGAAATTTCCAGATAT 876
Qy      198   Gly-----AlaLeuAlaSerLeuIleArgSerVal 207
Db      877  GGGAAAGTTTTCAGAGAAATTAAGTTAAATATGCCAGCTGCGAGGGCCAAAGAGATC 936
Qy      208   AlaSerPheSer-----IleYrSerProHisThrGlyIleGlnGluYrGln 223
Db      937  ATTCTTACTCCGACCTGCTGACTGACTTGTCTCT-----GGGGTGAAGTCTTATCCA 990
Qy      224   Asp----- 224
Db      991  GATGTTGGAATCTTCTCGAGGTGTGTCCAGCGTGAAATATCTAAATCTGAATGCT 1050
Qy      224  ----- 224
Db      1051  GCAGAGAACCTTTCACACCAAGTTTACCAGCAATGAAATGCTTATAGCGCTGAATT 1110
Qy      225  -----GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      1111  GCAGAGCGTGTGTCTTCCAAAGATTCCTGTTCAATTTGATATCTAATGACGACAG 1170
Qy      241   MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db      1171  AAGCTCTAGAA-----AAATGGGTGGC 1194
Qy      261   LysThrYrProAsp----- 265
Db      1195  TCAGACCAACCAAGATGAGCTGAGAGAGAAAGTCTCAAAAGTCCCTACATGTTGACCT 1254
Qy      265  ----- 265

```

```

Db      1255  GCGTTTACTGGAACCTTTTCTACCAAAAGTCAAGATGACATCTCACTACCAATGAA 1314
Qy      266  ---ThrAspSerPheAsnThrValAlaGluIleThrGlySerLysYrProGluGlnVal 284
Db      1315  GTGACAAAGATTTTACATGTGATAGTACTCTCAGAGAGACATGGAACCAACAGATAT 1374
Qy      285   ValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGly 304
Db      1375  GTATTCTGGAGAGTCAACCGGACTCATAGGCTGTT-----GGTGTATTTGACCCCTCAG 1428
Qy      305   GlyGlyAlaPheIleSerTrpGlu-----AlaLeuSerLeuIleLysAspLeuGly 321
Db      1429  AGTGAGACACTGTGTTCTCAATGAAATTTGTGAGAGAGCTTTGGAACACTGAAAGAAAGG 1488
Qy      322   LeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGluGlnGlyVal 341
Db      1489  TGGAGACTTGAAAGAAATTTGTTTGGCAAGCTGGATGCAAGAAATTTGGCTTCTT 1548
Qy      1549  GGTTCTACTGAGTGGGAGAG-----GAGATTCAGACACTCTTCAA 1590
Db      362   GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyVal 381
Qy      1591  GAGCGTGGCGTG---GCTTATATTAATGCTGACTCATCTATAGAGAAATCTACACTCTG 1647
Db      382   ArgAlaIleMetGluGluValMet---SerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Qy      1648  AGAGTGATTTGTACACCGCTGATGTACAGCTTGTCAC-----AACCTAACAAAGAG 1701
Db      401   Leu---SerHisGlyGluGlyYrAsp-IleAsnPheTrpIleGlnAlaGlyVal 417
Qy      1702  CTGAAAGCCCTGATGAGAGCTTTGAAAGCAATCTCTTATGAAAGTTGAGACTA 1756

RESULT 8
US-08-394-152A-1
; Sequence 1, Application US/08394152A
; Patent No. 5935818
GENERAL INFORMATION:
APPLICANT: Israel, Ron S.
APPLICANT: Heaton, Warren D.W.
APPLICANT: Falt, William R.
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM 330 466 DX2
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,152A
FILING DATE: 24-FEB-95
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41426-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2653 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Carcinoma
; IMMEDIATE SOURCE:
; CLONE: Prostate-Specific Membrane Antigen
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 262..2511
; US-08-394-152A-1

Alignment Scores:
Pred. No.: 3,036-09 Length: 2653
Score: 166.00 Matches: 105
Percent Similarity: 32.47% Conserved: 70
Best Local Similarity: 19.48% Mismatches: 169
Query Match: 6.87% Indels: 195
DB: 2 Gaps: 23

US-09-745-763-36 (1-472) x US-08-394-152A-1 (1-2653)

OY 3 PheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCyseSerGlyLysAla 22
DB 355 TTCTTCTCTCTCGGCTTCTCTCTGTTTATTAATCCATCCATGAAGACTACTAC 414
OY 23 IleCyLysPheGlyIleSerLysArgThrPheGluGluIleLysGluGluIleAlaSer 42
DB 415 ATTACTCCAAAGCATTAATGAAGACATTTTGTGATGTAATGAAGCTGAG----- 465
OY 43 CyseGlyAspValAlaLysAlaIleLeuLeuAlaValIleGlyLysAlaGlnAsnArg 62
DB 466 -----ACATCAAGAAAGTTCTTATATATTT----- 492
OY 63 SerTyrgLysArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGlySerLys 82
DB 493 -----ACACAGATTCACCATTTGACGAGAACAGA 522
OY 83 -----AsnLeuGlyLysAlaIleGlnIleMetTyrgLysAsnLeuGlnAsnArgLys 99
DB 523 CAAAACTTTCAGCTTCGCAAGCAAAATTCATCCAGTGGAAGAAATTTGGCTGATTTCT 582
OY 100 LeuGluLysValHisLeuGluProVal-----ArgIleProHisTrp 113
DB 583 GTTGAGCTAGACACATTATGATGTCCGTGTCTCTACCCAAATTAAGACTCATCTAC 642
OY 114 -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
DB 643 ATCTCAATTAATTAAGAGATGAAGATGATTTCAACACATCATTTATTTGAACCA--- 699
OY 127 IleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146
DB 700 -----CCTCTCTCCAGGAATAT 714
OY 147 ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnArgAlaSerGluAla 166
DB 715 GAAAAATGTTTGGATATGTACACCTTTCAGTCTTCTCTCTCCAGGAATG---CCA 771
OY 167 ArgGlyLysIleValValTyraAsnGlnProTyrlleAsnTySerArgThr----- 183
DB 772 GAGGCGCATCTAGTG-----TATGTTAATATAGCAGGAATGAACACTTC 816
OY 184 -----ValGlnTyraArgThrGlnGlyAlaValGluAlaAlaLysVal 197
DB 817 TTTAAATTCGAACGGACATGAATAATCAATTCCTCGGGAATAATGTATTCGAGATAT 876
OY 198 Gly-----AlaLeuAlaSerLeuIleArgSerVal 207
DB 877 GGGAAAGTTTTCAGAGAAATAGGTTAAATATGCCAGCTGGCGGGGCCCAAGAGAGTC 936
OY 208 AlaSerPheSer-----IleTyserProHisThrGlyIleGlnGluTyrgln 223

```

```

DB 937 ATTCTTACTCCGACCTGCTGACTACTTGTCTCT-----GGGGTGAAGCTCTATCA 990
OY 224 Asp----- 224
DB 991 GATGTTGGAATCTTCTCTGAGGTGTGTCCAGCGTGAATATCTTAATTCGATGCT 1050
OY 224 ----- 224
DB 1051 GCAGAGACCTCTCAGACCAAGTTACCGCAATATATGCTTATAGCGGTGAAT 1110
OY 225 -----GlyValProLysIleProThrAlaCySilThrValGluAspAlaGlu 240
DB 1111 GCAGAGCTGTGTGCTTCCAGATATCTTTCATCTCAATATGATATGATGACACAG 1170
OY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
DB 1171 AAGCTCTAGAA-----AAATGGGTGGC 1194
OY 261 LysThrTyrrProAsp----- 265
DB 1195 TCAGACACACCAAGATACAGCTGAGAGAGAGTCTCAAGTCCATATGTTGACCT 1254
OY 265 ----- 265
DB 1255 GCCTTACTGGAACCTTTTCTACACAAAAGTCAAGATGACATCCACTTACCAATGAA 1314
OY 266 ---ThrAspSerPheAsnThrValAlaGluIleThrLysSerLysTyrrProGluGlnVal 284
DB 1315 GTGACAAAGATTAATCAATGTGATGTGACTCTCAGAGAGAGTGAACCAAGACATAT 1374
OY 285 ValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGly 304
DB 1375 GTCATTCGGAGAGCTTCACCGGACTCATGGTGT-----GGTGGATTTGACCTCAG 1428
OY 305 GlyLysAlaPheIleSerTrpGlu-----AlaLeuSerLeuIleLysAspLeuGly 321
DB 1429 AGTGAGCAGAGCTGTGTTTCATGAATTTGTGAGAGCTTTGGAAACATGAAAGAGGG 1488
OY 322 LeuArgProLysArgThrPheLeuArgLeuValLeuThrPheArgLysGluGlyVal 341
DB 1489 TGGAGACCTGAAACAAACATTTTGTTCAGAGCTGGAGATGAGAAATTTGGTCTTCT 1548
OY 342 GlyAlaPheGlnTyrrTyrgLysLeuHisLysValAsnIleSerAsnTyrrSerLeuValMet 361
DB 1549 GGTCTTACTAGTGGCAGAG-----GAGATTCAGAGCTCTTCAA 1590
OY 362 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLysAla 381
DB 1591 GAGCGTGGCGTG---GCTATATTAATGCTGATCATCTATAGAAGGAAACTACACTCTG 1647
OY 382 ArgAlaIleMetGluGluValMet-----SerLeuLeuGlnProLeuAsnIleThrGlnVal 400
DB 1648 AGAGTGAATGTATACACCGCTGATGTACAGCTTGATGAC-----AACCTTAACAAAGAG 1701
OY 401 Leu---SerHisGlyGlyGlyThrAsp-IleAspPheTrpIleGlnAlaGlyVal 417
DB 1702 CTGAAAGCCCTGATGAAGCTTTGAAGCAAACTCTTATATGAAGTTGAGCTA 1756

RESULT 9
US-08-705-477B-1
; Sequence 1, Application US/08705477B
; Patent No. 6569432
; GENERAL INFORMATION:
; APPLICANT: Israel, Ron S
; APPLICANT: Heaton, Warren D.W.
; APPLICANT: Fair, William R.
; APPLICANT: Overfelli, Quathek
; APPLICANT: Pinto, John
; TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF
; FILE REFERENCE: 1769/41426-G
; CURRENT APPLICATION NUMBER: US/08/705, 477B
; CURRENT FILING DATE: 1996-08-29

```

```

NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 1
LENGTH: 2653
TYPE: DNA
ORGANISM: Homo sapiens
US-08-705-477E-1

Alignment Scores:
Pred. No.: 3,03e-09 Length: 2653
Score: 166.00 Matches: 105
Percent Similarity: 32.47% Conservative: 70
Best Local Similarity: 19.48% Mismatches: 169
Query Match: 6.87% Indels: 195
DB: 4 Gaps: 23

US-09-745-763-36 (1-472) x US-08-705-477E-1 (1-2653)

QY 3 PheLeuIlePheAlaPhePheGlyValHisLeuSerLeuCySerGlyLysAla 22
DB 355 TTCTTTCTCTCGGCTTCCTCTCGGCTTATAAATCTCCCAATGAAGCTACTAAC 414
QY 23 IleCyAlaSerGlyIleSerLeuArgThrPheGluGluIleGlyGluIleAlaSer 42
DB 415 ATTACTCCAAAGCATATATGAAAGCATTTTGGATGTAATGAAAGCTGAG----- 465
QY 43 CysGlyAspValAlaIleValAlaIleLeuLeuAlaValTyGlyLysAlaGlnAsnArg 62
DB 466 -----AACATCAAGAAAGTTCTTATATATTT----- 492
QY 63 SerTyGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGlySerLys 82
DB 493 -----ACACAGATACCAACATTTAGCAGGAACGAA 522
QY 83 -----AsnLeuGluLysAlaIleGlnIleMetTyGlnAsnLeuGlnAspGly 99
DB 523 CAAAACTTCAGCTTGCAAGCAAAATTCATCCCAAGTGAAGAAATTTGGCTGATTCCT 582
QY 100 LeuGluLysValHisLeuGluProVal-----ArgIleProHisTrp 113
DB 583 GTTAGCTAGACATTTATGATGTCTGTGTCTTACCAAAATTAAGACTCATCCCAACTAC 642
QY 114 -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
DB 643 ATTCATATATTAATGAAGATGAAGATGATTTCAACACATCATTTTGAACCA--- 699
QY 127 IleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146
DB 700 -----CCTCTCCAGGATAT 714
QY 147 ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnArgAlaSerGluAla 166
DB 715 GAAAAATGTTGGATATTGTACACACCTTCAGTCTTCTCTCTCCAGGAATG---CCA 771
QY 167 ArgGlyLysIleValValTyArgAsnGlnProTyIleAsnTySerArgThr----- 183
DB 772 GAGGGCGATCTAGTC-----TATCTTAATCTATGCGAAGTGAAGACTTC 816
QY 184 -----ValGlnTyArgThrGlnGlyAlaValAlaGluAlaLysVal 197
DB 817 TTTAAATGGAACGGACGATGAATCAATTCCTCTGGAAATTTGTAATGCCAGATAT 876
QY 198 Gly-----AlaLeuAlaSerLeuIleArgSerVal 207
DB 877 GGGAAAGTTTTCAGAGAAATTAAGTTAAATATGCCAGCTGGCGAGGGCCAAAGAGATC 936
QY 208 AlaSerPheSer-----IleTySerProHisThrGlyIleGlnGluTyGln 223
DB 937 ATTCTCTACTCGACCCCTGCTGACTACTTGTCTCT-----GGGGTAAGTCCATCA 990
QY 224 Asp----- 224
DB 991 GATGGTGGAAATCTTCTCTGAGGTGTGTCCAGCGTGAATATCTTAATCTGAATGAT 1050

```

```

QY 224 ----- 224
DB 1051 GCAGAGACCCCTCACACAGGTTACCCAGCAAAATGATATGCTTATAGCGTGAATT 1110
QY 225 -----GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
DB 1111 GCAGAGCTGTGTCTTCCAAAGTATTCCTGTTCATTCATTCATGATGATGACAG 1170
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
DB 1171 AAGCTCTAGAA-----AAAAAGGATGGC 1194
QY 261 LysThrTyProAsp----- 265
DB 1195 TCAGCACCCACAGATAGACACTGAGAGGAAGTCTCAAAAGTCCCTACATGTTGACCT 1254
QY 265 ----- 265
DB 1255 GGCTTTACTGAAACTTTTCTACAAAAAGTCAAGATGCATCCACTTACCAATGAA 1314
QY 266 -----ThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyProGluGlnVal 284
DB 1315 GTGACAAAGATTTATCAATGTGATAGTACTCTCAGAGAGACAGTGAACGACAGATAT 1374
QY 285 ValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGly 304
DB 1375 GTCAATTTGGAGAGTCAACCGGACTCATGSGTGT-----GTTGATATGACCTCAG 1428
QY 305 GlyValAlaPheIleSerTrpGlu-----AlaLeuSerLeuIleLysAspLeuGly 321
DB 1429 AGTGAAGACCTGTGTGTATGAAATGTGAGAGACTTGTGAACACTGAAAGAAAGAGG 1488
QY 322 LeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGlyVal 341
DB 1489 TGGAGACCTTGAAGAAATTTGTTTGAAGCTGGATGCCAAGAAATTTGCTCTTCT 1548
QY 342 GlyAlaPheGlnTyTyGlnLeuHisLysValAsnIleSerAsnTySerLeuValMet 361
DB 1549 GGTTCATCTAGAGTGGCAGAG-----GAGATTTCAAGACTCTTCA 1590
QY 362 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLysAla 381
DB 1591 GAGGCTGGCGG-----GCTTATATTAATGCTGACTCATCTATGAAAGAACTACACTCG 1647
QY 382 ArgAlaIleMetGluGluValMet-----SerLeuGlnProLeuAsnIleThrGlnVal 400
DB 1648 AGAGTTGATTGTACACCGCTGATGTACAGCTTGATAC-----AACCTAACAAAGAG 1701
QY 401 Leu-----SerHisGlyGluGlyThrAsp-IleAsnPheTrpIleGlnAlaGlyVal 417
DB 1702 CTGAAAGCCCTGATGAGAGGCTTGAAGCAAAATCTTATGAAAGTTGACTA 1756

RESULT 10
US-09-164-034B-1
GENERAL INFORMATION:
APPLICANT: Mincheff, Milcho S.
Loukinov, I. Dmitri
TITLE OF INVENTION: Immunotherapy of Cancer Through Expression
of Truncated Tumor- or Tumor-Associated Antigen
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: William S. Ramsey,
Ramsey, Cook, Looper & Kurlander, LLC
STREET: 10420 Little Patuxent Parkway, Suite 250
CITY: Columbia
STATE: Maryland
COUNTRY: USA
ZIP: 21044
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: PC

```

```

OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/164,034B
FILING DATE: 30-Sep-1998
ATTORNEY/AGENT INFORMATION:
NAME: Ramsey, William S.
REGISTRATION NUMBER: 32,715
REFERENCE/DOCKET NUMBER: br11
TELECOMMUNICATION INFORMATION:
TELEPHONE: (410) 992-9660
TELEFAX: (410) 992-9540
SEQUENCE DESCRIPTION: SEQ ID NO: 1
US-09-164-034B-1

Alignment Scores:
Pred. No.: 8.3e-09 Length: 2133
Score: 161.00 Matches: 93
Percent Similarity: 33.48% Conservative: 63
Best Local Similarity: 19.96% Mismatches: 139
Query Match: 6.66% Indels: 171
Gaps: 21

US-09-745-763-36 (1-472) x US-09-164-034B-1 (1-2133)
QY 76 ProAglSerGlySerLys-----AenLeuGluValAlaIleGlnIleMetLys 92
DB 115 CCACATTGACGACGACACGACAACTTCAGCTTCGACCAATTCATCCAGTGG 174
QY 93 GlnAenLeuGlnIleApsGlyLeuGluValIleLeuGluProVal----- 108
DB 175 AAGAAATTTGGCTGCTGTTGCTGAGTACACATTAATGATGCTGTTGCTTACCA 234
QY 109 -----ArgIleProHisLys-----GluArgGlyGluGlu-----Ser 119
DB 235 AATAAGACTCATCCCACTACATCTCAATTAATTAATGAATGAATGAAATTTTCAAC 294
QY 120 AlaValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIle 139
DB 295 ACATCATTAATTGAAACA----- 312
QY 140 GlyThrProProGluGlyIleThrAlaGluValLeuValIleThrSerPheApsGluLeu 159
DB 313 -----CCTCCGACGATATGAAATGTTTCGATGATTCACACCTTCAGTCTTTC 366
QY 160 GlnAgaGlnAseGlnIleArgGlyLysIleValIleValIleThrAsnGlnProTyrIleAsn 179
DB 367 TCTCTCTCAAGAGATG---CCAGAGGGCGATCTAGTG-----TATGTTAAC 408
QY 180 TyrSerArgThr-----ValGlnTyrArgThrGlnGly 190
DB 409 TATGACACGACGACGACGACCTTCTTAATTTGACGAGGACATGAAATCAATTCCTCTGG 468
QY 191 AlaValGlnAlaAlaLysValGly-----AlaLeu 200
DB 469 AAATTTGTAATGCGACATATGGAAGATTTTCAAGAGAAATGAAGTTAAATGAAGCCAG 528
QY 201 AlaSerLeuIleArgSerValAlaSerPheSer-----IleTyrSerProHis 216
DB 529 CTGGAGGGGGCCAAAGAGATCTTCTCACTCCGACCTGCTGACTACTTGTCTCTCT--- 585
QY 217 ThrGlyIleGlnGluTyrGlnAps----- 224
DB 586 ---GGGGTGAAGTCTTATTCAGATGCTTGGAATTTCTTCGAGGTGTGTCCAGCGGTGGA 642
QY 224 ----- 224
DB 643 AATATCTTAATCTGAATGTCGACAGACCTCTCACACGAGTTACCCAGCAATGA 702
QY 225 -----GlyValProLysIleProThrAlaCys 233
DB 703 TATGCTTATAGCGGTGAATTCAGAGGCGTGTGCTTCCTCCAAAGATTCTGTTTCATCCA 762

```

```

QY 234 IleThrValGlnAsePalaGluMetSerArgMetAlaSerHisGlyIleVal 253
DB 763 ATGGATATCATATGACACAGAACCTCTTGAA----- 795
QY 254 IleGlnLeuLysMetGlyAlaLysThrTyrProAps----- 265
DB 796 -----AAATGGGTGGCTGACGACACCAATGATGACGAGAGGAAGTCTGAAA 846
QY 265 ----- 265
DB 847 GTGCCCTCAATGTTGACCTGCTTACTGGAACCTTTCTACACAAAAAGTCAAGATG 906
QY 266 -----ThrApsSerPheAsnThrValAlaGluIleThrGly 277
DB 907 CACATCCACTCTACCAATGAAGTGAAGATTAATTAATGATGATGATGATGATGATGATG 966
QY 278 SerLysTyrProGluGlnValValLeuValSerGlyHisLeuApsSerTyrPheApsValGly 297
DB 967 GCAGTGGACACGACACATATGTCATTCGTGGAGGTGACCGGAGCTCATGGGTGTTT--- 1023
QY 298 GlnGlyAlaMetApsApsGlyGlyAlaPheIleSerTyrGlu-----AlaLeu 314
DB 1024 ---GGTGTATGACCTCTCAAGTGGAGCAGCTGTTTCAATGAATTTGAGGAGCTTT 1080
QY 315 SerLeuIleLysApsLeuGlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThr 334
DB 1081 GCACACTGAAAGAAAGAGGTGAGACCTTGAAGAACATTTGTTGCAAGCTGGGAT 1140
QY 335 AlaGluGlnGlnGlyValGlyAlaPheGlnTyrGlnLeuHisLysValAsnIle 354
DB 1141 GCACAAAGAAATTTGGCTCTTCTGTTCTTCACTGACAGTGGCCAGAG----- 1182
QY 355 SerAsnTyrSerLeuValMetGluSerApsAlaGlyThrPheLeuProThrGlyLeuGln 374
DB 1183 GAGATTCAGACCTCTTCAAGACGCTGGCGTG---CCTTATATTAATGCTGACTATCT 1239
QY 375 PheThrGlySerGluValAargAlaIleMetGluGluValMet-----SerLeuLeuGln 393
DB 1240 ATAGAGGAAACTACCTGAGAGTGTATGTTACACCGGATGTTGACAGCTTGTGTAAC 1299
QY 394 ProLeuAsnIleThrGlnValLeu---SerHisGlyGlnGlyThrAps-IleAsnPheTr 412
DB 1300 -----AACCTCAACAAAAGAGCTGAAGACCTGATGAGGCTTTGAAGCAATCTCT 1353
QY 412 PileGlnAlaGlyVal 417
DB 1354 TATGAAAGTTGACTA 1369

RESULT 11
US-08-705-477E-100
; Sequence 100, Application US/08705477E
; Patent No. 6569432
; GENERAL INFORMATION:
; APPLICANT: Israeli, Ron S.
; APPLICANT: Heaton, Warren D.W.
; APPLICANT: Fair, William R.
; APPLICANT: Overfield, Ouathex
; APPLICANT: Pinto, Don
; TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF
; FILE REFERENCE: 1769/41426-G
; CURRENT APPLICATION NUMBER: US/08/705,477E
; CURRENT FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 100
; LENGTH: 2387
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-705-477E-100

Alignment Scores:
Pred. No.: 9.99e-09 Length: 2387
Score: 161.00 Matches: 93

```

Percent Similarity: 33.48% Conservative: 63
 Best Local Similarity: 19.96% Mismatches: 139
 Query Match: 4 Gaps: 21
 DB: 4

US-09-745-763-36 (1-472) x US-08-705-477E-100 (1-2387)

```

Qy 76 ProArgLeuSerGlySerLys-----AsnLeuGluYsAlaIleGlnIleMetYr 92
Db 236 CCAATTTAGAGAGAGAGAGAGAACTTTCAGCTTGGAAGCAATTCATCCAGCG 295
Qy 93 GlnAsnLeuGlnIleArgLysLeuGluYsAlaIleGlnIleMetYr 108
Db 296 AAGAAATTTGGCTGAGATTCCTGAGCTAGCAATTCATTCCTGCTTCCACCA 355
Qy 109 -----ArgIleProHisTrp-----GluArgIleGluGlu-----Ser 119
Db 356 AATAGACTATCCCACTACATCTCAATTAATTAATGAAGATGAATGAGATTTTCAC 415
Qy 120 AlaValMetLeuGluProArgIleHisIleAlaIleLeuGlyLeuGlySerSerIle 139
Db 416 ACATCATTTATTGAACA-----433
Qy 140 GlyThrProProGluGlyIleThrAlaGluValValThrSerPheAspGluLeu 159
Db 434 -----CCTCTCCAGATATGAATAATGTTCCGATATTTACACCTTTCAGTCTTC 487
Qy 160 GlnArgArgAlaSerGluAlaArgGlyYsIleValValIleValIleAsn 179
Db 488 TCTCTCAAGGAATG-----CCAGAGGGCGATCTAGTG-----TATGTAAAC 529
Qy 180 TySerArgThr-----ValGlnIleYrArgThrGlnGly 190
Db 530 TATGCAGGAATGGAAGCTTTTAATGGAAGGAGCATGAATAATGATGCTTGG 589
Qy 191 AlaValGluAlaIleValGly-----AlaLeu 200
Db 590 AAAATTTGTAATGGCAGATATGGAGAAAGTTTTCAGAGGAATTAAGTTAAATGCCAG 649
Qy 201 AlaSerLeuIleArgSerValAlaSerPheSer-----IleYrSerProHis 216
Db 650 CTGGCAGGGGCCAAGAGATCTTCTACCTCCAGCCTGTGACTACTTGTCTCT-- 706
Qy 217 ThrGlyIleGlnGluYrGlnAsp----- 224
Db 707 ---GGGGTGAAGCTCTATCCAGATGTTGGAATCTTCTGAGAGGTGCTCCAGCGTGA 763
Qy 224 ----- 224
Db 764 AATATCTTAATCTGAATGTGACAGAGACCTCTCACACAGGTTACCCAGAAATGA 823
Qy 225 -----GlyValProLysIleProThrAlaCys 233
Db 824 TATGCTTATAGCGGTGAATTCAGAGAGCTGTGCTCTCCAGATATTCCTGTTCAATCA 883
Qy 234 IleThrValGluAspAlaGluMetMetSerArgMetAlaSerHisGlyIleValIleVal 253
Db 884 ATTGATATCTATGATGACAGAGAGCTCTAGAA-----916
Qy 254 IleGlnLeuYsMetGlyAlaIleYsThrYrProAsp----- 265
Db 917 -----AAAATGGGTGGCTCAGACACACAGATACAGCTGAGAGAGAAAGTCTAAA 967
Qy 265 ----- 265
Db 968 CTGGCCCTAAGATGTTGACCTGGCTTTAGTGAACCTTTCTACACAAAATCAAGATG 1027
Qy 266 -----ThrAspSerPheAsnTrpValAlaGluIleThrGly 277
Db 1028 CACATCCAGCTTAGCAATGAAGTACAGAAATTTACATGATGATGATGATCTCAGAGGA 1087
Qy 278 SerYrYrProGluGlnValValIleValSerGlyHisLeuAspSerTrpAspValGly 297
Db 297 -----

```

Db 1088 GCAGTGGACACAGACATATGATTCCTGGAGGTCACCGGACATCATGCTGTTT--- 1144
 Qy 298 GlnGlyAlaMetAspAspGlyGlyValAlaPheIleSerTrpGlu-----AlaLeu 314
 Db 1145 ---GGTGTATTTGACCTCTCAGAGTGGAGAGCTGTTGTTTCAATGAATTTGAGAGAGCTTT 1201
 Qy 315 SerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThr 334
 Db 1202 GGAACACTGAAAAGAGAGGTGGAGACCTGAGAACATTTGTTTGGACCTGGAT 1261
 Qy 335 AlaGluGluGlnGlyValGlyValAlaPheGlnTrpYrGlnLeuHisIleValAsnIle 354
 Db 1262 GCAGAAATTTGCTCTTCTGTTCTTCTGATCTACAGAGCGGACAG-----1303
 Qy 355 SerAsnTrpSerLeuValMetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGln 374
 Db 1304 GAGATTCAGACCTCTTCAAGAGCGTGGCTG---CCTATATTAATGCTGACTATCT 1360
 Qy 375 PheThrGlySerGluYsAlaArgAlaIleMetGluGluValMet---SerLeuGln 393
 Db 1361 ATGAGAGAAACTACACTGAGCTGATGATTGACACCGCTGATGACCTTGTACAC 1420
 Qy 394 ProLeuAsnIleThrGlnValLeu---SerHisGlyGluGlyThrAsp-IleAsnPheTr 412
 Db 1421 -----AACCTTAACAAAGAGCTGAAGAGCCTGATGAGAGCTTTGAAGCAATCTCTT 1474
 Qy 412 PheGlnAlaGlyVal 417
 Db 1475 TATGAAGTTGACTA 1490

RESULT 12
 US-09-252-991A-1872
 : Sequence 1872, Application US/09252991A
 : Patent No. 6551795
 : GENERAL INFORMATION:
 : APPLICANT: Marc J. Rubenfield et al.
 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 : FILE REFERENCE: 107196, 136
 : CURRENT APPLICATION NUMBER: US/09/252,991A
 : PRIOR FILING DATE: 1999-02-18
 : PRIOR APPLICATION NUMBER: US 60/074,788
 : PRIOR FILING DATE: 1998-07-27
 : NUMBER OF SEQ ID NOS: 3142
 : SEQ ID NO 1872
 : LENGTH: 765
 : TYPE: DNA
 : ORGANISM: Pseudomonas aeruginosa
 : US-09-252-991A-1872

Alignment Scores:
 Pred. No.: 4,1e-08 Length: 765
 Score: 149.00 Matches: 65
 Percent Similarity: 40.87% Conservative: 38
 Best Local Similarity: 25.79% Mismatches: 95
 Query Match: 4 Gaps: 11
 DB: 4

US-09-745-763-36 (1-472) x US-09-252-991A-1872 (1-765)

```

Qy 146 IleThrAlaGluValIleValIle-----ThrSer 155
Db 83 GTCACGGCCAAAGTGTCCGGTGAACCTTCTCGGCGCCGCAACACTCCACAGC 142
Qy 156 PheAspGluLeuGlnArgAlaSerGluAlaArgGlyYsIleValIleValYrAsnGln 175
Db 143 GTTTCAGAGCGGAGAAAGCTTCCCACTTCCCGCGCGCTGATGATCCAGCGC 202
Qy 176 ProYrIleAsnTrpSerArgTrpValGlnIleYrArgThrGlnGlyAlaValGluAla 195
Db 203 GCACACTGCAACTTCGAG-----CAGAAAGCGGAGAAAGCGCGC 241

```



```

QY 196 lysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerPro 215
Db 242 GCGCGCGCGCGCGCGCGCGGTGATCATCTTCACCAAGGCGAAC-----ACCAAGCAC 2292
QY 216 HisThrGlyIle-----GlnGluTyrGlnIleAspGlyValProLysIle 229
Db 229 CGCAAGGCGCGTGAAGAAGCGACCGTGGCGGAGTCTCAACAGAGGCGGCATCCGGGTATC 355
QY 230 ProThrAlaCysAlleThrValGluAspAlaGluMetCysSerArgMetAlaSerHisGly 2494
Db 353 -----TTCGCCACTTACGACCAACGGCGTGGCGTGTGGCGAGACCCCGACTTCAG 4030
QY 250 IleIysIleValIleGlnLeuLysMetGlyAlaIleThrTyrProAspThrAspSerPhe 2695
Db 404 TTGCACCTGTGTGTCCACGTG-----GTACCGAAGAAGACCGACACTTAC 4488
QY 270 AsnThrValAlaGluIleThrGlySerLysTyrProGluIleValLeuValSerGly 2885
Db 449 AACGTGTGCGCGAG--ACCGTGGCGGACCCGAAACAAGTGTGATGTGTGGCGG 5095
QY 290 HisLeuAspSerTrpAspValGlyGlnGlyAlaMetCysAspGlyGlyGlyAlaPheIle 3093
Db 506 CACCTCGACTCGGTGTTCGAAGGCCCGCGGTATCAACACACAGCTTGGGACAGCGCC 5656
QY 310 SerTrpGlu---AlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeu 3282
Db 566 CACATGGAATGCGCCGTCTGCTGGCCAAAGCGCTG-----CCGTCAACAAGTGTG 6166
QY 329 ArgLeuValLeuTrpThrAlaGluGluGlnGlyGlyValGlyAlaPheGlnTyrTrGln 3488
Db 617 CGCTTCGCTGTGGTGGGGCGCGCGAAGACCGGCTGTGGTGGCTGCACCCACTACGTGCG 6767
QY 349 LeuHisLeuValAsnIleSer-----AsnTyr 3575
Db 677 -----AACCTGCGCCGGAAGAGAAGAAAGTCAAGCGCTTCACTGAACCTTC 7242
QY 358 SerLeuValMetGluSerAspAlaGlyThrPheLeu 369
Db 725 GACATCATCGCGCTCGCGCAACTTCGGCACTTCACTC 760

RESULT 13
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FIEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-2007.00
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 0.0615 Length: 4403765
Score: 149.00 Matches: 97
Percent Similarity: 34.73% Conservative: 60
Best Local Similarity: 21.46% Mismatches: 181
Query Match: 6.16% Indels: 114

```

DB: 3 Page: 21

US-09-745-763-36 (1-472) x US-09-103-840A-2 (1-4403765)

QY 75 GYPRoATgLeuSerGlySerLySAsnLeuGluValAlaIleGlnIleMetTyrgInAsn 94

Db 505153 GGCACtCGCGGGGTGGGCAcCCCTGGATATCAAGGCACGGCTGCATATGTGGTAAACACA 505212

QY 95 LeuGlnInAspGlyLeuGlnLySValHisLeuGluProValArgIleProHisTrpGlu 114

Db 505213 CTGCCCAACAGCGGTTTGAT-----GTCCAAACCCCGGAGTTCTCC 505254

QY 115 -----ArgGlyGluGluSerAlaValIleMetLeuGluProArgIleHisLyS 130

Db 505255 GCTCGCGTTTCAGGCGCAAAAGAGGGGTGGTGGACCTTCGGCGC-----AACACCGTg 505308

QY 131 AlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIleThrAlaGluVal 150

Db 505309 GAGCGGAAGGGCGCTCGAGTACAGCCTTCGGCACACCGCGGAGGGGTGAAGGGCCGGCTg 505368

QY 151 LeuValValThrSerPheAspGluLeuGlnArgAlaSerGlu----- 165

Db 505369 GTGGCTGCCCGCCGCGGACGACAGTCCGGGCTGCAGTCCGTGGACTACAGACAGGCTGCCg 505428

QY 166 AlaArgGlyLySIlleValValTyrgAsnGlnProTyrlleAsnTySerArgThrValGln 185

Db 505429 GTGTCCGGTGGCGGTGTCTGGTAATCGCGCGCTGTGCTTTGGC----- 505476

QY 186 TyrgArgThrGlnGlyAlaValGluAlaAlaLySValGlyAlaLeuAlaSerLeuIleArg 205

Db 505477 -----CAGAGGAAGACCGCACCGCCGCGGAGCGGGTGGCGCTGCATCTGCT 505527

QY 206 -----SerValAlaSerPheSerIleTySerProHisThrGlyIleGlnGlu 221

Db 505528 GACACATCGACGAGACGGCGATGGGCGGACCTCGGGGGCTAATACGACGTC----- 505581

QY 222 TyrgInAspGlyValProLySleIleProThrAlaCysIlleThrValGluAspAlaGluMet 241

Db 505582 -----AAGATCCCGGTGTGAGTGTCCACAG----- 505608

QY 242 MetSerArgMetAlaSerHisGlyIleLySIlleValIleGlnLeuLyS----- 257

Db 505609 -----TCGGTCGATTCACGTTACGACGACGCGACAGTCTGGg 505641

QY 258 ---MetGlyAlaLySThrTyrgProAspThrAspSerPhe-----AsnThrValAla 273

Db 505642 CCAACACCGCTCAAGCTACCGCGGACGACCCAAAGTTTCAGAGCCCGCAACGTCATCGCG 505701

QY 274 GluIle---ThrGlySerLySThrTyrgProGluGlnValValLeuValSerGlyHisLeuAsp 292

Db 505702 CAGACGAAGACGGGTGTCTGCG-----GCCAACGTGGTATGACAGGTGCGGATTTGGAC 505755

QY 293 SerTrpAspValGlyGlnGlyAlaMetAspAspAspGlyGlyAlaPheIleSerTrpGlu 312

Db 505756 AGCGTTCCGGAAGGACCGGACATCAACGACACGGCTCGGAGTGGCT-----GCG 505806

QY 313 AlaLeuSerLeuIleLySAspLeuGlyLeuArgPro-----LySArgThrLeuArgLeu 330

Db 505807 GTTTCGAAAGCGGACGTGAGCTGGGGAACTCACCGGCATGTGTCCAAACCGGTAACGGTTC 505866

QY 331 ValLeuTrpThrAlaGluGlnGlyValGlyValAlaPheGlnIleTyrgIleLeuHis 350

Db 505867 GCCTTCTCGGGGCGCGAGGAATTCGGCGCTGATGGGTGACGAACACTACGTGCGAGTGTG 505926

QY 351 LySValAsn-----IleSerAsnTySerLeuValMetGluSer 363

Db 505927 GACATCGACGCGCTCAAAAGCATCGCGCTGATCTGAACTTGACACATGTGTGCGCTGCCG 505986

QY 364 AspAlaGlyThrPheLeuProThrGly-----LeuGlnIleThrGlySer 378

Db 505987 AACCCGGGTACTTACCTACGATGACGATGACAGGTGCTGCCGTAGACCGCGCGGTACG 506046

QY 379 -----GluLySAlaArgAlaIleMetGluGluValMetSerLeuLeuGlnPro 394

Db 506047 CCGGTGTCGCCGGAAGGCTCGGCGCGGTATCGAGCGACCTTCGTCCTCATCTGAAGATG 506106
 Qy 395 LeuAniIeThr--GlnValIeuSerHisGlyGluGlyThrAspIleAsnPhetPile 413
 Db 506107 GCCGCGAAGACCGCGCAGAGACACCTTCGTTCGACGCGTGTCCGACTACAGACGCGCTTCAG 506166
 Qy 414 GlnIaGlyValProGlyAlaSerIleu----- 422
 Db 506167 CTGGCGGGATTCCTTCCTCGGCGGTGCGCTTCCTCCGCGCTGAGGTCAAGATCCGCGAG 506226
 Qy 423 -----LeuAspAspLeuTyrIleTyrPhe 430
 Db 506227 CAAGCGGACGCTCTGGGGCGCGACCGCGCGACGACCTTCCTTCGATCCCAACTAT----- 5062777
 Qy 431 PhePheHisHisSerHisGlyAspThrMetThrValMetAspProIleGlnMet----- 448
 Db 506278 -----CACCGAAGACAGACACCTTCGAGACATATGACCGGACCGCGCTCGGTATC 506328
 Qy 449 AsnValAlaAlaAlaValIleTAlaValAlaSerTyr 460
 Db 506329 AACGGCGCTGGCGGTCCGTACCGCGGTGGGTGTAT 506364
 RESULT 14
 US-09-103-840A-1
 Sequence 1, Application US/09103840A
 Patent No. 6294328
 GENERAL INFORMATION:
 APPLICANT: FLEISCHMAN, Robert D.
 APPLICANT: WHITE, Owen R.
 APPLICANT: FRASER, Claire M.
 APPLICANT: VENTER, John C.
 TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 TITLE OF INVENTION: TUBERCULOSIS
 FILE REFERENCE: 24366-20007.00
 CURRENT APPLICATION NUMBER: US/09/103,840A
 CURRENT FILING DATE: 1998-06-24
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 4411529
 TYPE: DNA
 ORGANISM: Mycobacterium tuberculosis
 OTHER INFORMATION: H37Rv
 US-09-103-840A-1
 Alignment Scores:
 Pred. No.: 0.0617 Length: 4411529
 Score: 149.00 Matches: 97
 Percent Similarity: 34.73% Conservative: 60
 Best Local Similarity: 21.46% Mismatches: 181
 Query Match: 6.16% Indels: 114
 DB: 3 Gaps: 21
 US-09-745-763-36 (1-472) x US-09-103-840A-1 (1-4411529)
 Qy 75 GlyProArgLeuSerGlySerIysAsnLeuGluValAlaIleGlnIleMetTyrGlnAsn 94
 Db 503710 GGCACCTCGCGCGGTGGGACCCCTGCTATCAGCGACGGGTGCACATATGCGTAACACA 503765
 Qy 95 LeuGlnIaAspGlyLeuGluIlyValHisLeuGluProValArgIleProHisTyrPheGlu 114
 Db 503770 CTGCCGACACAGCGGTTTGTAT-----GTGCAAAACCCCGGAGTTCTCC 503811
 Qy 115 -----ArgGlyGluGluSerAlaValMetLeuGluProArgIleHisIlyPhe 130
 Db 503812 GCTCCGCGTTCAGAGCCCAAAAAGGGGTGTACCTCGGCGCG-----AACACCGTG 503866
 Qy 131 AlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluIlyIleThrAlaGluVal 150
 Db 503866 GAGCGAAGGGCGCTCGAGTACAGCTTCGCGACACCGCGGACGGGATGACGGCGCGCTG 503922
 Qy 151 LeuValValIleThrSerPheAspGluLeuGlnIaArgAlaSerIle----- 165

Db	503926	GTGGCTGCCCCCGGACGACAGTCGCGGCTGCAGTCCGTCGGATTCAGACAGCGCTGCCG		503985
Qy	166	AlaArgglyLysIleValValTyrAsnGlnProTyrIleAsnTyrSerArgThrValGln		185
Db	503986	GTGTCGGGTGGCGGTGCTCTGGTAGATCGCGGCGCTGTCGCTTTGGC-----		504033
Qy	186	TyrArgThrGlnGlyAlaValGlnAlaIleValLysValGlyAlaLeuAlaSerLeuIleArg		205
Db	504034	-----CAGAAAGAAAGACGACGCGCGGACGCGCGGTGGCGGTGATCTATGCT		504084
Qy	206	-----SerValAlaSerPheSerIleTyrSerProHisThrGlyIleGlnGlu		221
Db	504085	GACAAATCGACGACGACGAGCCGATGGCGCGCACCTTGGGGGCTTAATATCCAGCTC-----		504138
Qy	222	TyrGlnArgGlyValProLysIleProThrAlaCysIleThrValGlnAlaGlnMet		241
Db	504139	-----AAGATCCGGGTGGGTGAGTGTACCAAG-----		504165
Qy	242	MetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLys-----		257
Db	504166	-----TCGGTCGATTCACAGCTACCGGACAGCTCTGGG		504198
Qy	258	--MetGlyAlaLysThrTyrProAspThrAspSerPhe-----AsnThrValAla		273
Db	504199	CCAAACACCGTCACAGCTACGCGCGACGCCAACAAAGTTTTCAGAGCCCGCAACGTATCGCG		504258
Qy	274	GluIle---ThrGlySerLysTyrProGlnGlnValIleValSerGlyHisLeuAsp		292
Db	504259	CAGACGAAAGACGGGGTCTGTCG-----GCCAACGTGGTATGCGACGTCGCGATTTGCAC		504312
Qy	293	SerTyrAspValGlyGlnGlyAlaMetAspArgGlyGlyValAlaPheIleSerTyrGlu		312
Db	504313	AGCGTTCGGAAGGACCCGCGCATCACACACACGCTCGGGAGTGACT-----GCG		504363
Qy	313	AlaLeuSerLeuIleLysAspLeuGlyLeuArgPro-----LysArgThrLeuArgLeu		330
Db	504364	GTTCTGGAAACGCGCAGTCGACGTGGGAACTCACCCGCATGTCTCAACCGCGTACGGTTC		504423
Qy	331	ValLeuTyrThrAlaGlnGlnGlnGlyValGlyAlaPheGlnTyrTyrGlnLeuHis		350
Db	504424	GCGTTCGTGGGGCGGAGGAATTCGCGCATGTGGGTGTCAGAACTACGTGACGTGCTG		504483
Qy	351	LysValAla-----IleSerAsnTyrSerLeuValMetGluSer		363
Db	504484	GACATCGACGCGCTCAAAAGCGATCGCGCTGTATCTGAACCTTGACATGTTGGCGTCCGCG		504543
Qy	364	AspAlaGlyThrPheLeuProThrGly-----LeuGlnPheThrGlySer		378
Db	504544	AACCGGGTACTTCACTTACGACGCGTACACAGTCGCTGCGCTGACGCGCGGGTCCAG		504603
Qy	379	-----GlnLysAlaArgAlaIleMetGlnGlnValMetSerLeuLeuGlnPro		394
Db	504604	CCGGTGTGTCGCGGAAGGCGCTCGGCGCGGTATCGAGCGCACCTTGCTGCCATATCGAAGATG		504653
Qy	395	LeuAsnIleThr---GlnValLeuSerHisGlyGlnGlyThrAspIleAsnPheTyrIle		413
Db	504654	GCCGGCAAGACCGGCGGACACCTCTGTTGACGCGTGGTCCGACTACGACGCGCTTCACG		504723
Qy	414	GlnAlaGlyValProGlyAlaSerLeu-----		422
Db	504724	CTGGCGGGTATCTCTTGGGGTGGCTGTCTTCCGCGGCTGAGGTCAAGAGTCCGCGGAG		504783
Qy	423	-----LeuAspAspLeuTyrTyrSerPhe		430
Db	504784	CAAGCCGAGCTCTGGGGGGGCGACCGCCGACGAGCGCTTTCGATCCCAACTAT-----		504834
Qy	431	PhePheHisHisSerHisGlyAspThrMetThrValMetAspProLysGlnMet-----		448
Db	504835	-----CACCAAGACGACACCTCTGACCAATATGACCGGACCGCGCTCGGTATTC		504885
Qy	449	AsnValAlaIleAlaValThrAlaValAlaSerTyr		460

Db 50486 AACGGCGCTGCGCGCGGTCGCGGTTGTAT 504921

RESULT 15

US-09-016-434-1352

Sequence 1352, Application US/09016434

Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young

APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

TITLE OF INVENTION: PATHWAY GENE EXPRESSION

NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,434

FILING DATE: HEREMITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1352:

SEQUENCE CHARACTERISTICS:

LENGTH: 4080 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: 937432

US-09-016-434-1352

Alignment Scores:

Pred. No.: 0.000263 Length: 4080

Score: 127.00 Matches: 88

Percent Similarity: 32.96% Conservative: 59

Best Local Similarity: 19.73% Mismatches: 187

Query Match: 5.25% Indels: 113

DB: 4 Gaps: 18

US-09-745-763-36 (1-472) x US-09-016-434-1352 (1-4080)

QY 76 ProATGLeuSerGlySerLySasnleuGluYvalaIleGlnIleMetYrGlnasnleu 95

Db CCTCGTAGGCGTGCATCTCAAAAAGATGAATAATCTTGGCTGTATCTTGAATAATCAATT 782

QY 96 GlnGlnaerGlyLeuGlnYvalaIhIleuGlnPro-----ValArgIleProhIstYr 113

Db CGTGAATTTAAACTCAGCAAAAGCTCGCGGTGATCAAACTTTGTTAAGATTCAGGTCAA 842

QY 114 GlnaGlyGlnGlnSerAlaValMetleuGlnProaGlnIleHIsIleAlaIleu 133

Db GACAGCGCTCAAAACTCGGTGATCATAGTTGATAAAGAGGTAAGACTTGTAACTGGTG 902

QY 134 ---GlyLeuGlySerSerIleGlyThrProProGlnGlyIleThrAlaGlnValleuVal 152

Db GAGAACTCGGGGTATATGTCGCTATAGTAAGGCTGCAACAGTTACTGTAACTGGTC 962

QY 153 ValThrSerPhe-----AspGlnleuGlnArgArgAlaSerGlnAlaArgIlyLys 169

Db CATCGTAATTTTGGTACTAAAAAAGATTTGAGATTTTAAACCTCGTGAATGATCT 1022

QY 170 IleValValTyrAengInProTyrIleasnTyrSerArgThrValaGlnTyrArgThrGln 189

Db ATAGTATGTCAGAGCAGGAAATATCACTTTCAGAAAAGTT----- 1067

QY 190 GlnAlaValGlnAlaAlaIlySvalGlyAlaIleuAlaSerleuIle---ArgSerValAla 208

Db -----GCAATGCTGAAGCTTAAATGCAATGTGTGTGATATATCATGACCAGACT 1121

QY 209 SerPheSerIle----- 212

Db AAATTTCCCATTTGTTAAACGAACTTTCATTTTGGACATCTGGGACAGGT 1181

QY 213 -----TyrSerPro-----HisThrGlyIle---GlnGlnTyrGln 223

Db GACCTTACACACCTGATTCCTTCCTCAATCACTCACTGTTCCACCATCTCGGTCA 1241

QY 224 AspGlyValProIlySIIleProThrAlaCySIIleThrValaGlnaPalaGlnMetMet--- 242

Db TCAGATTCCTTATATATACCTGTCAGACAATCTCCAGAGCTGCTGCAGAAAAGCTGTTT 1301

QY 243 -----SerArgMetAla 246

Db GGAATATGAGAGAGACTGCTCCTGACTGAAAACAGACTTACATGTAGATGTA 1361

QY 247 -----SerHisGlyIleIlySIIleValIleGlnleuIlySmetGlyAlaIlyThrTyr 263

Db ACCTCAGAAACAGAAATGTCACACTCAGTCTG-----AGCAATGTCTG 1406

QY 264 ProaPThrAerSerPheasnThrValaIleGlnIleThrGlySerIlySyrProGln 283

Db AAAGAGATAAAAATTTTACATCTTGAATTTTAAAGCTTTGTAGAACAGATCAC 1466

QY 284 ValValleuValSerGlyHisIleuAerSerThrAerValaGlyGlnGlyAlaMetAerAer 303

Db TATGTTAGTTGGGGCCAGAGAGATCATG-----GGCCCTGAGCTGCAAAATCC 1520

QY 304 GlyGlyGlyAlaPheIleSerTrpGlnAlaIleuSerleuIlySAsPleu----- 320

Db GGTGAGGACACAGCTCTCTTATGAACCTGCCAGATGTTCTCAGATATGCTTTAAA 1580

QY 321 ---GlyLeuArgProIlySArgThrLeuArgLeuValleuTrpThrAlaGlnGlnGly 339

Db GATGGTTTACGCCACAGACAGCATTATCTTCCAGTTGGAGTGTGAGACTTGA 1640

QY 340 GlyValaGly-----AlaPhe 344

Db TCGGTTGTGTCAGTGAATGCTAGAGGATACCTTTCCTCCATTTAAAGCTTTC 1700

QY 345 GlnTyrTyrGlnleuHIsIlyS-----ValasnIleSerAsn----- 356

Db ACTTATATTAATCTCGATTAAGCGGTTCTGTGACCAAGCACTTCAAGGTTTCCGAC 1760

QY 357 -----TyrSerleuValMetGlnSerAspAlaGlyThrPheIlyProThrGlyLeu 373

Db CCACGTTGTATACGTTATGAGAAAACATGCAAAATGAGACATCCGGTATACGGG 1820

QY 374 GlnPheThrGlySerGlnIlySAlaArgAlaIleMetGlnIlyValMetSerleuLeuGln 393

Db CAATTTCTATATTCAGACAGACAGACAGTGGCCAGCAAACTTGAACATCTTAA----- 1874

QY 394 ProleuasnIleThrGlnValleuSerHisGlyGlnGlyThrAerIleasnPheTrpIle 413

Db 1875 -----GACATGCTGCTTCCCTTCTTGA 1901

QY 414 GlnAlaGlyValProGlyAlaSerleuLeu-----AspAerleuTyrIlySyrPhe 430

Db	1902	TATTCTGGAAATCCAGAGATTCTTCTGTTTTCGAGACAGATTATCTTATTG	1961
Qy	431	PhePheHisHisSerHis	436
		:::	
Db	1962	GG-TACCACCATGGACAC	1978

Search completed: December 23, 2003, 02:25:48
Job time : 2472 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 22, 2003, 23:44:51 / Search time 408 Seconds
(without alignments)
3854.902 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417

Sequence: 1 MKELIFAFGFGVHLLSLCSG.....AVNAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delop 6.0, Delext 7.0

Searched: 221978 segs, 1666101734 residues

Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/cg2_1/USPTO.spool_p/US09745763/runat.22122003.113318.25304/app_query.fasta.1.647
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rmpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09745763 @CGN_1.1.85 @runat.22122003.113318.25304
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*

1: /cg2_6/ptodata/2/pubna/US07_PUBCOMB.seq:*

2: /cg2_6/ptodata/2/pubna/PT_NEW_PUB.seq:*

3: /cg2_6/ptodata/2/pubna/US06_NEW_PUB.seq:*

4: /cg2_6/ptodata/2/pubna/US06_PUBCOMB.seq:*

5: /cg2_6/ptodata/2/pubna/US07_NEW_PUB.seq:*

6: /cg2_6/ptodata/2/pubna/US07_PUBCOMB.seq:*

7: /cg2_6/ptodata/2/pubna/US08_NEW_PUB.seq:*

8: /cg2_6/ptodata/2/pubna/US08_PUBCOMB.seq:*

9: /cg2_6/ptodata/2/pubna/US09A_PUBCOMB.seq:*

10: /cg2_6/ptodata/2/pubna/US09B_PUBCOMB.seq:*

11: /cg2_6/ptodata/2/pubna/US09C_PUBCOMB.seq:*

12: /cg2_6/ptodata/2/pubna/US09_NEW_PUB.seq:*

13: /cg2_6/ptodata/2/pubna/US09_NEW_PUB.seq:*

14: /cg2_6/ptodata/2/pubna/US10A_PUBCOMB.seq:*

15: /cg2_6/ptodata/2/pubna/US10B_PUBCOMB.seq:*

16: /cg2_6/ptodata/2/pubna/US10_NEW_PUB.seq:*

17: /cg2_6/ptodata/2/pubna/US60_NEW_PUB.seq:*

18: /cg2_6/ptodata/2/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match Length	ID	Description
------------	--------------------	----	-------------

1	2417	100.0	1851	9	US-09-745-763-35	Sequence 35, Appl
2	2417	100.0	1863	11	US-09-984-271-28	Sequence 28, Appl
3	2116	87.5	1778	10	US-09-917-800A-505	Sequence 505, Appl
4	1338	55.4	1134	11	US-09-984-271-95	Sequence 95, Appl
5	508.5	21.0	357	10	US-09-833-381-1929	Sequence 1929, Ap
6	443	18.3	427	10	US-09-833-381-1930	Sequence 2337, Ap
7	190	7.9	1346	10	US-10-325-430-14	Sequence 13, Appl
8	188	7.8	3152	13	US-10-325-430-13	Sequence 9, Appl
9	188	7.8	2653	13	US-10-352-157-9	Sequence 13, Appl
10	169	7.0	2653	10	US-09-869-708-536	Sequence 536, App
11	166	6.9	2653	13	US-09-873-319-451	Sequence 451, App
12	166	6.9	2653	13	US-09-960-706-716	Sequence 716, App
13	166	6.9	2653	13	US-10-210-120-5	Sequence 5, Appl
14	166	6.9	2653	13	US-10-341-434-70	Sequence 70, Appl
15	166	6.9	2653	13	US-10-117-937-7	Sequence 7, Appl
16	166	6.9	2653	15	US-10-094-699-2	Sequence 131, App
17	166	6.9	2653	15	US-09-978-295A-617	Sequence 617, App
18	164	6.8	2558	10	US-09-978-697-617	Sequence 617, App
19	164	6.8	2558	10	US-09-978-832A-617	Sequence 617, App
20	164	6.8	2558	10	US-09-978-832A-617	Sequence 617, App
21	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
22	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
23	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
24	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
25	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
26	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
27	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
28	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
29	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
30	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
31	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
32	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
33	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
34	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
35	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
36	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
37	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
38	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
39	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
40	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
41	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
42	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
43	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
44	164	6.8	2558	11	US-09-978-832A-617	Sequence 617, App
45	164	6.8	2558	13	US-10-143-030A-617	Sequence 617, App

ALIGNMENTS

RESULT 1
US-09-745-763-35
Sequence 35, Application US/09745763
Patent No. US20020065394A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallie, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Merberg, David
Treacy, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM

NUMBER OF SEQUENCES: 219

CORRESPONDENCE ADDRESS:

ADDRESS: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/745,763
  FILING DATE: 18-Jun-2000
  CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
  NAME: Sprunger, Suzanne A.
  REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (617) 498-8284
  TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 35:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1851 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-745-763-35

Alignment Scores:
  Pred. No.: 3,756-286      Length: 1851
  Score: 2417.00           Matches: 472
  Percent Similarity: 100.00%   Conservative: 0
  Best Local Similarity: 100.00%   Mismatches: 0
  Query Match: 100.00%         Indels: 0
  DB: 9                      Gaps: 0

US-09-745-763-36 (1-472) x US-09-745-763-35 (1-1851)
QY 1 MetLysPheLeuIlePheAlaPheGlyValHisLeuLeuSerLeuCySerGly 20
DB 99 ATGAATTCCTTATCTTCGATTTTCGGTGCTGTCACCTTTTATCCCTGCTGCG 158
QY 21 LysAlaIleCysValAsnGlyIleSerLysArgThrPheGluGluIleValGly 40
DB 159 AAGCTATATGCAAGATGCGATCTCTTAAGAGCACTTTGAGAAATTAAGAAATA 218
QY 41 AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAla 60
DB 219 GCCAGCTGAGAGATGTTGCTTAAGCAATCATCAACCTGATGTTATGTAAGCCCG 278
QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSer 80
DB 279 AACGATCTTATGACGATTTGGCAGCTTCTGTTGATCTGTTGAGCCGAGCTGAG 338
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100
DB 339 TCCAGAACCTTAGAAGAACCCATCCAAATTAATGACCAAAACCTGAGCAGATGGCTG 398
QY 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSer 120
DB 399 GAGAAAGTTCACTGAGGACCACTGAGAAATACCCCACTGGGAGAGGAGAAATCAG 458
QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIle 140
DB 459 GTGATGCTGAGACCAAGATTTCAATGATGACATCTGGGTCTTGAGCAGCATTTGG 518
QY 141 ThrProArgGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluGln 160
DB 519 ACTCTCCAGAAAGCAATTACAGCAGAGTTCTGGGTGATCACTCTTTCATGAACTGC 578
QY 161 ArgAlaGlaserGluAlaArgGlyLysIleValIleTyrAsnGlnProTyrIleAsn 180
DB 579 AGAAGGCTCAGAGCAAGAGGAGGAGATGTTGTTTAAACAACCTTACATCACTAC 638
QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaIleLysValGlyLeu 200
DB 639 TCAAGGACGCTGATATACGAGACGAGGGGCGGTGAGAACTGCCAAGGTGGGCTTGG 698

```

```

QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
DB 699 GCATCTCTCATTCATCCGTCGCTCTCTCCATCTACAGTCTCTCACACAGGATTCAG 758
QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
DB 759 GAATACCGAGATGCGTGGCCCAAAATTCAAACACCTGATTAAGGTGAGAGCAGAA 818
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
DB 819 ATGATGTCAGAAATGCTTCTCATGGATCAAAATTCATTCAGCTAAAGATGGGGCA 878
QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
DB 879 AAGACCTACCAAGTATCATGATCTTCAACACTGATACAGAGATCATCTGGAGCAATAT 938
QY 281 ProGluGlnValIleLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300
DB 939 CCAGAACAGGTGTGACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 998
QY 301 MetAspAspGlyGlyValAlaPheIleSerThrGluAlaLeuSerLeuIleLysAspLeu 320
DB 999 ATGATGATGCGGTGAGCTTATATCATGAGACACTCTCATTAATTAAGATCTT 1058
QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrPheAlaGluGluGlnGlyGly 340
DB 1059 GGGCTGCTCCAAAGAGAGACTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
DB 1119 GTTGCTGCTTCCAGTATATCATGATGATGATGATGATGATGATGATGATGATGATG 1178
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnIleThrGlySerGlyLys 380
DB 1179 ATGAGATCTGACGAGAGAACTTCTTACCCACTGGGCTGCAATTCACCTGAGGAG 1238
QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
DB 1239 GCCAGGCGCATCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1298
QY 401 LeuSerHisGlyGluGlyThrAspIleAsnThrPheIleGlnAlaGlyValProGlyAla 420
DB 1299 CTGAGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1358
QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
DB 1359 AGCTACTGATGACTTATACAGATATTTCTTCTTCAATCACTCCACGAGAGACCAT 1418
QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValIleThrAlaValIleSerTyr 460
DB 1419 ACTGTCATGATCCAAAGCAGATGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTG 1478
QY 461 ValValAlaAspMetGluMetLeuProArgSer 472
DB 1479 GTTGTCAGACATGAGAAATGCTGCTGATGCTC 1514

RESULT 2
US-09-984-271-28
; Sequence 28, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984, 271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482, 273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092, 921
; PRIOR FILING DATE: 1998-07-15

```


PRIOR APPLICATION NUMBER: 60/092,922
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/092,956
PRIOR FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 1863
TYPE: DNA
ORGANISM: Homo sapiens
US-09-984-271-28

Alignment Scores:
Pred. No.: 3,79e-286 Length: 1863
Score: 2417.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-745-763-36 (1-472) x US-09-984-271-28 (1-1863)

QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20
DB 99 ATGAAATTCCTTATCTTCGCAATTTTCGCTGCTTCACTTTATCTGCTGCTGCG 158
QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
DB 159 AAACCTATATCAAGAAATGCGATCTCTAAGAGACTTTTGAAGAAATTAAGAAATA 218
QY 41 AlAserCySGlyAspValAlaIleLeuAsnLeuAlaValTYRGLYValAlaGln 60
DB 219 GCCAGCTGTGAGATGTTGCTAAGCAATCATCACTAGCTGTTTATGTAAGCCAG 278
QY 61 AsnArgSerTYRGLYValGluLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80
DB 279 AACGATCTCTATGAGCGATGCGACTTCTGCTGTTGACTGTGAGCCAGACTAGTGGC 338
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTYRGLYAsnLeuGlnAspGlyLeu 100
DB 339 TCCAAGAACTTGAAGAAAGCAATTCATTAATGTAACCAAACTGCGAGCAAGATGGGCTG 398
QY 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluGluSerAla 120
DB 399 GAGAAAGTTCACTGAGAGCCAGTGAATACCCACTGCGAGAGGAGAAAGATCACT 458
QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLysSerSerIleGly 140
DB 459 GTGATGTGAGCAAGAAATTCATTAAGATACCATCTGCGGCTTGGCAGCAGCATTTGG 518
QY 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
DB 519 ACTCTCTCAAGAGGATTACAGCAAGATTTCTGGGTGGAGCTCTTTCAGTAAGCTGAG 578
QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTYRAsnGlnProTYRILEAsnTYR 180
DB 579 AGAAGGCTCTAGAAAGCAAGAGGAAAGTTGTTGTTTAAACCAACTTAATCACTAC 638
QY 181 SerArgThrValGlnTYRArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200
DB 639 TCAAGAGCGGTGCATATACCAAGCGAGGGGCGGTGAAGCTGCCAAGTTGGGGCTTGG 698
QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTYRSerProHisThrGlyIleGln 220
DB 699 GCATCTCTCATTCGATTCGCTGCGCTCTTCTTCATCTCACTCACTCAACAGGATTCAG 758
QY 221 GluTYRGLYAspGlyValProLysIleProThrAlaCysIleThrValGlnAspAlaGlu 240
DB 759 GAATTCAGAGATGGCGTCCCAAGATTCACACAGCTTATTAACGTGGAAGATGCAGAA 818
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
DB 819 ATGATGTCAAGAAATGCTTCTCATGTGATCAAAATTTTCATTCAGCTAAAGATGGGGCA 878

QY 261 LysThrTYRProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTYR 280
DB 879 AAGACCTACCCAGATCTGATTTCTTCAACACTGTAGCAAGATCACTGGAGCAAAATAT 938
QY 281 ProGluGlnValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
DB 939 CCAGAACAGGTTGACTGTCAGTGCACATCTGCAGCTGGAGTGGATGTTGGGAGGGTGGC 998
QY 301 MetAspAspGlyGlyValAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
DB 999 ATGATGATGCGCGTGAAGCTTTATATCATGGAAGCACTCTCATTTAATAAGACTT 1058
QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGlyGly 340
DB 1059 GGGCTGTGTCACAAAGAGACTCTGCGCTGTCTCTGAGCTGCAGAAAGAAAGAGTGA 1118
QY 341 ValGlyAlaPheGlnTYRGLYValLeuHisLysValAsnIleSerAsnTYRSerLeuVal 360
DB 1119 GTTGTGCTCTTCCAGTATATCATGTTACACAGGTAAATATTTCCAACTACAGCTGCTG 1178
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
DB 1179 ATGAGCTGACCGCAGAACCTTCTTACCCTGCGCTGCAATTCATCTGCGACGTGAAG 1238
QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
DB 1239 GCCAGGCCATCATGAGAGGTTATGAGCTGTGCAAGCCCTCAATATCACTCAAGTTC 1298
QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
DB 1299 CTGAGCCATGAGAGAGGACAGACATCACTTTTGATCCAACTGAGAGCTGGAGCC 1358
QY 421 SerLeuLeuAspLeuTYRGLYTYRPhenPheHisHisSerHisGlyAspThrMet 440
DB 1359 AGTCTACTTATATCATTAACAATATTTCTTTCATATCACTCCACGAGACACCATG 1418
QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleTrpAlaValIleSerTYR 460
DB 1419 ACTGTCAATGATCCAAAGCAAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTTAT 1478
QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
DB 1479 GTTGTGCACACATGAGAAAGATGCTGCTAGTCC 1514
RESULT 3
US-09-917-800A-505
Sequence 505, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Caselle, Arthur
APPLICANT: Blaschoff, Michael
APPLICANT: Gene Logic, Inc.
TITLE OR INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/230,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/230,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457

? PRIOR FILING DATE: 2001-06-13
 ? PRIOR APPLICATION NUMBER: US 60/298,884
 ? PRIOR FILING DATE: 2001-06-19
 ? PRIOR APPLICATION NUMBER: US 60/303,459
 ? PRIOR FILING DATE: 2001-07-09
 ? NUMBER OF SEQ ID NOS: 1740
 ? SOFTWARE: PatentIn Ver. 2.1
 ? SEQ ID NO 505
 ? LENGTH: 1778
 ? TYPE: DNA
 ? ORGANISM: *Rattus norvegicus*
 ? FEATURE:
 ? OTHER INFORMATION: Genbank Accession No. US20020119462A1 AF097722
 ? US-09-917-800A-505

ORGANISM: Homo sapiens
US-09-984-271-95

Alignment Scores:
Pred. No.: 5,99e-154 Length: 1134
Score: 1338.00 Matches: 261
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 55.36% Indels: 1
DB: 11 Gaps: 0

US-09-745-763-36 (1-472) x US-09-984-271-95 (1-1134)

```

QY 211 SerIleTyrSerProH1stHrGlyIleGlnGlnIleTyrGlnAspGlyValProIleS1lePro
Db 1 TCcATCTACAGCTCCTCACAGAGTATTCAGAAATACAGAGATGGCGGCCCAAGATTCa 60
QY 231 ThrAlaCysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGlyIle 250
Db 61 ACAGCCTGTATTACGTGGAGAGATGACAGAAATGATGTCAGATGGCTTCTCATGGATC 120
QY 251 LysIleValIleGlnIleuLeuMetGlyValAlaYstHrTyrProAspThrAspSerPheAsn 270
Db 121 AAAATTGTCACTACGCTAAAGATGGGGGCAAAAGCTTACCCAGATCTGATTCCTTCAAC 180
QY 271 ThrValAlaGluIleThrGlySerIleTyrProGluGlnValIleuValSerGlyHis 290
Db 181 ACTGTACAGAGATCACTGGAGAGCAAAATATCCAGAACGGTTGACTGCTCAGTGGACAT 240
QY 291 LeuAspSerTyrAspValGlyGlnGlyValaMetAspAspGlyGlyValaPheIleSer 310
Db 241 CTGACACAGCTGGAGTGGGAGAGGTCCAGATGATGAGCGGTGGAGCCTTTATATCA 300
QY 311 TrpGluAlaLeuSerIleuIleuAspLeuGlyLeuAspGlyPheGlyThrLeuArgLeu 330
Db 301 TGGGAACACCTCTCACTATTAAAGATCTTGGCGCTGCCCAAGAGAGACTCGCGGTG 360
QY 331 ValLeuTrpThrAlaGluGlnGlnGlyValaGlyValaPheGlnIleTyrTrpGlnLeuHis 350
Db 361 GTGCTCTGGAGCTGAGAAAGAACAGAGTGGAGTGGCTTCCAGTATTATCAGTTAC 420
QY 351 LysValAsnIleSerAsnTyrSerLeuValMetGluSerAspAlaGlyThrPheLeuPro 370
Db 421 AAGGTAAATATTTCACACTACAGTCTGTGATGAGTGTGACGAGAACTTCTTACC 480
QY 371 ThrGlyLeuGlnPheThrGlySerGlyValaAlaArgAlaIleMetGluIleValMetSer 390
Db 481 ACTGGGCTGCATTCACACTGGCAGTGAAGGCGCAGGCGC-ATCATGAGAGAGTTATGAGC 539
QY 391 LeuLeuGlnProLeuAsnIleThrGlnValLeuSerHisGlyGlyGlyThrAspIleAsn 410
Db 540 CTGCTGGAGCCCTCAATATACCTCAGTCTGAGCCATGAGAGAGGAGACACCTCAAC 599
QY 411 PheTrpIleGlnAlaGlyValaProGlyValaSerLeuLeuAspAspLeuTyrLysTyrPhe 430
Db 600 TTTTGGATCCAAAGCTGAGAGTCCCTGAGCCAGTCTACTGATGACTTATACAGATTTC 659
QY 431 PhePheHisHisSerHisGlyAspThrMetThrValMetAspProLysGlnMetAsnVal 450
Db 660 TTCTTTCATCACTCCCAAGAGACCCATGACTGATGATCCAAAGACAGATGATATT 719
QY 451 AlaAlaAlaValaValaValaValaSerTyrValaValaAlaAspMetGluGlnMetLeuPro 470
Db 720 GCTGTGCTGTTGGGCTGTTGTTCTTATGTTGTTGTCAGACATGAGAAATGCTGCT 779
QY 471 ArgSer 472
Db 780 AGGTCC 785

```

RESULT 5
US-09-833-381-1929
Sequence 1929, Application US/09833381
Patent No. US20020132090A1

GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1929
LENGTH: 357
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(357)
OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1929

Alignment Scores:
Pred. No.: 7.58e-53 Length: 357
Score: 508.50 Matches: 98
Percent Similarity: 92.37% Conservative: 11
Best Local Similarity: 83.05% Mismatches: 7
Query Match: 21.04% Indels: 2
DB: 10 Gaps: 1

US-09-745-763-36 (1-472) x US-09-833-381-1929 (1-357)

```

QY 321 GlyLeuArgProIlyAspArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGly 340
Db 7 GGCCTCCGA---AAGAGAGCTCTGCGCTGCTCTGAGCTGAGAGAAAGAGAGCA 63
QY 341 ValGlyAlaPheGlnIleTyrGlnLeuHisGlyValaAsnIleSerAsnTyrSerLeuVal 360
Db 64 ATGGTCCCTCCCGATTTATGAGCTACATAGGCAAAATTTCCAGTACAGCTTGGTG 123
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
Db 124 ATGGAAGCTGACTCAGAGAACCTTCTTACCACTGAGTGCAGTTCAGTGCAGTGCAG 183
QY 381 AlaArgAlaIleMetGluGlnValaMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db 184 GCCAGGCTATCATGAAGAGAGATCATATCTTCTGCAACCCCTCAATGTCACCAAGCTC 243
QY 401 LeuSerHisGlyGlnGlyThrAspIleAsnPheTrpIleGlnAla-GlyValProGlyAl 420
Db 244 TTTAGTAAATGAGAGAGAGATGATTTAATTCTTGATCCAGCTTGGAGTNCCTGGAGC 303
QY 420 AspLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGly 437
Db 304 CAGCTGCNTGATGATGATTTGATCAAGTATTCCTTTTCCATCATTCACACGGA 355

```

RESULT 6
US-09-833-381-1930
Sequence 1930, Application US/09833381
Patent No. US20020132090A1
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1930
LENGTH: 427
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: misc feature
LOCATION: (1)-(427)
OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1930

Alignment Scores:
Pred. No.: 1,12e-44 Length: 427
Score: 443.00 Matches: 91
Percent Similarity: 85.84% Conservative: 6
Best Local Similarity: 80.53% Mismatches: 14
Query Match: 18.33% Indels: 2
DB: 10 Gaps: 1

US-09-745-763-36 (1-472) x US-09-833-381-1930 (1-427)

QY 1 MetlysrpHeuLlePhealaphneGlyValHIsleuSerLeuCyserGly 20
DB 91 ATGAGGTCTCTTTCTTCTGTTCAAT-----GTTCACTTTTAGCCTGGCTGCT 144
QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleGluGlu 40
DB 145 AAGCTGATTCAGAAATGCTGTTCTCAGCAACATTTCCGAAATTAAGAAATTA 204
QY 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
DB 205 GCCAATGAAAGATGTTGCTTAAGCAATTAACAACCTTGCTTTATGTTAAATACAG 264
QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
DB 265 AACCGCTCTAAGACGTTTGAGACTTCTAGTTGATAGCTTGAGACCAAGACTGAGTGC 324
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnGlnAspGlyLeu 100
DB 325 TCTAAGAACTTGAAGAACTTATCAATCATATGTTCCAAACCTGACAGATGGCTTA 384
QY 101 GluLysValHisLeuGluProValArgIleProHisTyr 113
DB 385 GAAATGTTCACCTGAGAGAGTCAANAATACCCCACTGG 423

RESULT 7

US-09-974-300-2337
Sequence 2337, Application US/0974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
TITLE OF INVENTION: Methods for Monitoring Multiple Gene
FILE REFERENCE: 10085,500-US
CURRENT APPLICATION NUMBER: US/09/974,300
PRIOR FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680,598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 2337
LENGTH: 1346
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-2337

Alignment Scores:
Pred. No.: 8.09e-13 Length: 1346
Score: 190.00 Matches: 110
Percent Similarity: 40.28% Conservative: 89
Best Local Similarity: 22.27% Mismatches: 215
Query Match: 7.86% Indels: 80
DB: 10 Gaps: 23

US-09-745-763-36 (1-472) x US-09-974-300-2337 (1-1346)

QY 2 LysPheLeuIlePheAlaPhePheGlyValHIsleuSerLeuCyserGlyLys 21
DB 10 AAATGATGATGTTCCGATTC-----GCGTATGATCATTTGACGGGGCC 54
QY 22 AlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleGluIleAla 41
DB 55 GTGCTCCCGCATGCA-----ACGGGAATGCACCTCAGCGGCCCTCAGAAACAGCC 108
QY 42 SerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAla----- 59
DB 109 ATCCGAAAAGATATCGAAAA-----TTACGCAAAAAATTCAC 147
QY 60 GlnAsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSer 79
DB 148 GAGAACCGCGCTATCAAGACATTTACATTTAAGCAAAAGATGGCGCGGTGACA 207
QY 80 GlySerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGly 99
DB 208 GGCACGCGCAGAAAGAAAAAGAGCGCGCTTTATCCCTCAGACAGATGAAAAAGTCAAT 267
QY 100 LeuGluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluSer 119
DB 268 CTG---AAAGTACACACAAAGTCAAGCTTACCTGACCGGCTGAAAG----- 315
QY 120 AlaValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIle 139
DB 316 -----ACGCTTACCTTACGGAATATAGCTGCTCGCGCGCGCGCGCTGCGCC 369
QY 140 GlyThrProProGluGlyIleThrAlaGluValLeu-----ValValThr 154
DB 370 CCGACTCAGCAGAAAGCTTGCGCTCTCTATGATGCGCGCTCGCGCTCGCGCC 429
QY 155 SerPheAspGluLeuGlnArgArgAlaSerGlnAlaArgGlyLysIleValValTyrAsn 174
DB 430 GACTTCACAGAA-----GAACGAAAGGAAATGCGCCGCT----- 465
QY 175 GlnProTyrIleAsnTyrSerArgThrValGlnTyrArgThrGlnGlyAlaValGluAla 194
DB 466 -----ATTGTGAAGAGAGAGCTGATTCATGAAAGCGCAAAACGCTGCGACCA 519
QY 195 AlaLysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTyrSer 214
DB 520 GCGCCAGCGGAGTATC-----ATTATATATACGTCGACAGCTCGCTCGTCACT 573
QY 215 ProHisThrGlyIleGlnGluTyrGlnAspGlyAlaProLysIleProThrAlaCysIle 234
DB 574 CCGATCTTACCGGTATTAAGTCGAT-----ATTCAGATTTGCGCGCTC 618
QY 235 ThrValGlnAspAlaGluMetSerArgMetAlaSerHisGlyIleLysIleValIle 254
DB 619 AAAAAGAGAGCGGGGAA-----AAGCTGCTTCTGAAACAAGAAGCATC----- 663
QY 255 GlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPheAsnThrValAlaGlu 274
DB 664 ---CTGAAGCTGAAGAGCTCATAAATCAAACTCGCAAAACGTAAATCGCGCTCGCAA 720
QY 275 IleThrGlySerIleTyrProGluGlnValValLeuValSerGlyHisLeuAspSerTyr 294
DB 721 GCAAAAGGTGCAAAATCTCGAC---ATCGTTATATGACTTCGCAATATGACAGCTT 777
QY 295 AspValGlyGlnGlyAlaMetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeu 314
DB 778 CCAATGCAACCGGAGGCAATGACAAATGCTCCGCACTTCAAGTCGCTTGAATCGGCC 837
QY 315 SerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThr 334
DB 838 CGAATTTTGAAGACA-----GTCCAGCGCAACAAGAAATTCGCTTATATACATTCGA 891
QY 335 AlaGluGluGlnGlyValGlyAlaPheGlnTyrTyr-----GlnLeu 349
DB 892 GCCGAAAGAAATCGAGCTCTCGGTTGCGCATTAATGACGACATTCGAATCAATCGAA 951
QY 350 HisLysValAsnIleSerAsnTyrSerLeuValMet-----GluSerAsp 364

```

Db      952 GTCAACGAGCGTTGGCACTTAATGATGTCGCGACAGCTGGAGAAATGCT 1011
Qy      365 AlaIyhrPheLeuProThr-----GlyLeuGlnPheThrGlySerGlu 379
Db      1012 TCACAGCTGATCAATATACCGCGACGCTGGCAAACTGCTGCTGACCTAAGTAA 1071
Qy      380 LysAlaArg--AlaIleMetGluValMetSerLeuGlnProLeuAsnIleThr 398
Db      1072 GCCGCTCTTAAACCTTGAAGAAAGACCTATTATTTTACATCAAGGCGATCATCCGAC 1111
Qy      399 GlnValLeuSerHisGlyGlyThrAsp--IleAsnPhe--TrpIleGlnAlaGly 416
Db      1132 CATGTCCTCATCATGAGACCGCGATCGACTCGGCAACTTATTGAGAGAG----- 1185
Qy      417 ValProGlyAlaSerLeuLeuAspAspLeuYrlyrThrPhePheHisSerHis 436
Db      1186 ---CCGGACAGCGCGCTTTGGAGCTTGATC-----CACACCCCT 1224
Qy      437 GlyAspThrMetThrValMetAspProGlyGlnMetAsnValAlaAlaValTrpAla 456
Db      1225 TACACACGATTAACATATCAGCAAGACAGCTGAAACAGCCGCTCAATATCGCGGA 1284
Qy      457 ValValSerTyrrValAlaAlaAspMetGluGlnMetLeuPro 470
Db      1285 ACAGCTGTGTATTACTTCACCAAGAGGAAACAGAAACCG 1326

```

RESULT 8

```

US-10-325-430-14
; Sequence 14, Application US/10325430
; Publication No. US2003015325A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Bech
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; FILE REFERENCE: MPI01-294PIRM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2223)
; US-10-325-430-14

```

Alignment Scores:

```

Pred. No.: 3,13e-12 Length: 2223
Score: 188.00 Matches: 112
Percent Similarity: 35.24% Conservative: 85
Best Local Similarity: 20.04% Mismatches: 186
Query Match: 7.78% Indels: 176
DB: 13 Gaps: 27

```

US-09-745-763-36 (1-472) x US-10-325-430-14 (1-2223)

```

Qy      58 LysAlaGlnAen-----ArgSerTyrrGluArgLeuAlaLeuLeuValAsp 72
Db      166 AAAGCTGAAACATCAATCATTTCTTGCTTTTACAAAGCTT----- 210
Qy      73 ThrValGlyProGlnLeuSerGlySerIlySerIlyAsn-----LeuGlnValAlaIleGln 89
Db      211 -----CTCATCTGCGACGAGACAGAAACAAATTTCTTGCTTGCCAGAAATCCAA 261
Qy      90 IleMetTyrrGlnAenLeuGlnIleAsnGlyLeuGlnValIleLeuGlnPro---Val 108

```

```

Db      262 ACCAGAGGAAGAAATTTGAGCTAGATTCAGCAAGTTGGTTCATATATGATGTCCTTA 321
Qy      109 ArgIleProHisTrpGluArgGlyGlnGluSerAlaValMetLeuGlnProAlaIleHis 128
Db      322 TCTTACCCCAATAGACAAATGCCCAACTATATATGATTTGATGAGAA-----CAT 372
Qy      129 LysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGlnGlyIleThrAla 148
Db      373 GAAACTGAGATTTTCAAAACA---TCATACCTTGAACCAACACACAGATGCTATGAGAT 429
Qy      149 GluValLeuValValThrSerPheAspGluLeuGlnAlaArgAlaSerGluAlaArgGly 168
Db      430 GTTACAAATATTTGGCCACCATATATATGCTTTCAGCCCAAGCGATG---CCAGAGGA 486
Qy      169 LysIleValValValTrpAsnGlnProTyrrIleAsnTyrrSerIlyThrValGlnTyrr 186
Db      487 GATCTTGTA-----TATGTAACTATGCTGCACTGAAGACTTTTTCAAA 531
Qy      187 -----ArgThrGlnGlyAlaValAlaGluAlaAlaLysValGlyAla 199
Db      532 CTGAAAGAGAGATGGGATCACTGATCTGAGAAAGTTGTTATGCAAGTATGAGTAA 591
Qy      200 Leu-----AlaSerLeuIleArgSerValAlaSerPheSerIle 212
Db      592 ATCTTCAGAGGAATAAATTAATTAATGATGATGATGATGATGATGATGATGATGATG 651
Qy      213 TyrrSer----- 214
Db      652 TACTCAGATCCAGCTGATCTTCTGCTGAGGTACAGCCATATCCCAAGATGAT 711
Qy      215 ---ProHisThrGlyIleGln----- 220
Db      712 CTTCCTGAAACTGACGCGCAGAGAGAAATGTGTTAATTTGAATGCTGCTGACCCA 771
Qy      221 -----GluTyrrGlnAsp 224
Db      772 CTCACCTCCAGGCTATCCAGCAAAAGATATACCTTCAGACTTGATGTTGAAGAGAGTG 831
Qy      225 GlyValProLysIleProThrAlaCysIleThrValAluAspAlaGlnMetSerArg 244
Db      832 GGAATCCCCCGAATACCTGTATCATCCCATGATGATATATATGACAGAAATATATTAAT 891
Qy      245 MetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAlaLysThr----- 262
Db      892 TACTTGGA---GGAATGCTCCACAGATTAAGAGTTGAAAGAGGCCCTTAATGTGAT 948
Qy      263 -----TyrrProAspThrAspSerPhe----- 269
Db      949 TATAGTATCGGACCTGCTTACAGGAGATGATTTCTTTCAGAAAGTTAGATGCATGTT 1008
Qy      270 -----AsnThrValAlaGluIleThrGlySerLys 279
Db      1009 TATAACATCAATAAATTTACAGAGATTATACATGATGATGATGATGATGATGATGATG 1068
Qy      280 TyrProGlnGlnValValLeuValSerGlyHisIleuAspSerTrpAspValGlyGlnGly 299
Db      1069 GAACCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1122
Qy      300 AlaMetAspAspGlyGlyGlyAlaPheIleSerTrpGlu-----AlaLeuSerLeu 316
Db      1123 GCTATTTGACCAACCAACAGATGGGCTGCTGCTTTTGCAAAATAATGCGCGAGATTTTGAAA 1182
Qy      317 IleLysAspLeuGlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGlu 336
Db      1183 CTGATGTAGTAAAGCTGAGAGCTTGAAGAAACTATATTTTTCACAGATGATGATGATG 1242
Qy      337 GluGlnGlyGlyValGlyAlaPheGlnTyrrGlnLeuHisLysValAlaMetIleSerAsn 356
Db      1243 GAAATTTGACTTCTGGGTTCCAGAGATGGGCTGAG-----GAGAAAT 1284
Qy      357 TyrrSerLeuValMetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThr 376

```

```

Db      1285 GTCAAAATCTCCAGAGAGACAT---GCTATATCACTCGATTCACTATAGAA 1341
Qy      377 GlycerGluYsAlaArgAla-----lleMeGluGluValMetSerleu 391
Db      1342 GGCATTTACTCTCAGAGTTAGCTACTCTCCCTCTTTACCAATTAGTATTAAGTCTG 1401
Qy      392 LeuGlnProLeu----- 395
Db      1402 ACAAAGAGATCCCGACCTGATGAGTTGAGAGTAATCACTGTATGAAGCTGG 1461
Qy      396 -----AenllethrglnValLeuSerHisGlyLeu 405
Db      1462 TTGAAAAAGACCTTCACTGAAATAATAATTTGCCATGAATGAAGCTGGAGTCT 1521
Qy      406 GlyThrAspIleAsnPhetripIleGln---AlaGlyValProGlyAla----- 420
Db      1522 GGAAGTGTCTTGAAGCTTATTTTCAAGACTTGAATATGCTTCAGGAGACCCGTTAC 1581
Qy      421 -----SerLeuLeuAspLeuTyrlTyrlPhePheHisSerHisGlyAsp 438
Db      1582 ACTAAGATTAAGAAAAACAGATAGTACAGACCTACCCAGTGTACACAACTTTATGAG 1641
Qy      439 ThrMetThrValMet-----AspPro-----LysGlnMetAsnValAla 451
Db      1642 ACATTTGATTTGTGAGAAATTTTATGACCCCACTTTAAACAACTTTCTGTGCT 1701
Qy      452 AlaAlaValTTPAlaValAlaSerTyrlValAlaAlaAspMetGluGluMetLeuPro 470
Db      1702 CAATTACGAGGACCTGGTA---TATGAGCTTGTGAT---TCTAAATCACTTCT 1752

RESULT 9
US-10-325-430-13
; Sequence 13, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Siles-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; FILE REFERENCE: MP101-294P1RNM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3152
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-325-430-13

Alignment Scores:
Pred. No.: 5,436-12 Length: 3152
Score: 188.00 Matches: 112
Percent Similarity: 35.24% Conservative: 85
Best Local Similarity: 20.04% Mismatches: 186
Query Match: 7.78% Indels: 176
Gaps: 27

US-09-745-763-36 (1-472) x US-10-325-430-13 (1-3152)
Qy      58 LysAlaGlnAsn-----ArgSerTyrlGluArgLeuAlaLeuLeuValAsp 72
Db      211 AAAGCTGAAGAAACATCAATCATTTCTGCTCTTTTACCAAGCTT----- 255
Qy      73 ThrValGlyProGlnLeuSerGlySerLysAsn-----LeuGluLysAlaIleGln 89
Db      256 -----CCCTCATCTGGCAGAACAGAAACAAATTTCTTGCTTCCCAAGAAATCCAA 306
Qy      90 lleMetTyrlGlnAsnLeuGlnIleAspGlyLeuGluLysValIleLeuGlnPro---Val 108

```

```

Db      307 ACCAGTGAAGAAATTTGACCTAGATTCAAGCCAAAGTTGGTTATATATGTCCTCTTA 366
Qy      109 ArgIleProHisTrpGluArgGlyGluGlnSerAlaValMetLeuGluProArgIleHis 128
Db      367 TCTTACCCCAATGAGACAAATGCCCACTATATATGATTTGGATGAGAA-----CAT 417
Qy      129 LysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIleThrAla 148
Db      418 GAAACTGAGATTTCAAAACA---TCATPACTTGAACCCACCAAGCTGCTATGAGAAAT 474
Qy      149 GluValLeuValIleThrSerPheAspGlyLeuGlnArgAlaSerGluAlaArgGly 168
Db      475 GTTCAAAATATTTGGCCACCATATATATGCTTTTCAGCCCAAGGCAAG---CCAGAGGA 531
Qy      169 LysIleValValTyrlAsnGlnProTyrlIleAsnTyrlSerArgThrValGlnTyrl 186
Db      532 GATCTTGTA-----TATGTGAACATATCTGCGACCTGAAGACTTTTTCAAA 576
Qy      187 -----ArgThrGlnGlyValAlaGluAlaIleValAlaIleValAla 199
Db      577 CTAGAAAGAGATGGGATCACTGATCTGGAAGATTGTATATGCAAGATATGAAAA 636
Qy      200 Leu-----AlaSerLeuIleArgSerValAlaSerPheSerIle 212
Db      637 ATCTTCAAGAGAAATAAAGTTAAATATGCAATGTATACAGAGACCATAGAAATCATCTTG 696
Qy      213 TyrlSer----- 214
Db      697 TACTCATCATCAGCTGACTACTTGTCTGTAGAGTACAGCATATATCCAAAGATGAAAT 756
Qy      215 ---ProHisThrGlyIleGln----- 220
Db      757 CTTCCTGGAAGTCAAGCCCAAGAGAGAAATGTATTAATTTGAATGTCGCTGTGACCCA 816
Qy      221 -----GluTyrlGlnAsp 224
Db      817 CTCACCTCAGGCTATCCAGCAAAAGAAATPACCTTTCAGACTGTGATGTTGAAGAAGAGTGT 876
Qy      225 GlyValProLysIleProThrAlaCysIleThrValAlaLysAspAlaGluMetMetSerArg 244
Db      877 GGAATCCCGGCAAAATCTGTACATCCCATGTGATATATATGACAGAAATATTTTTCGCG 936
Qy      245 MetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAlaLysThr----- 262
Db      937 TACTTGGGA---GGAATTGTCTCCACCAATGAAGTTGGAAGGAGCCCTTAATGTAGT 993
Qy      263 -----TyrProAspThrAspSerPhe----- 269
Db      994 TATAGTATCGGACCTGGCTTTACAGGAGATGATTTCTTCAAGAGTTAGATGATCATGTT 1053
Qy      270 -----AsnThrValAlaGluIleThrGlySerLys 279
Db      1054 TATAACATCATATAAATTACAGAGATTTCACAAATGTAGTTGGAACTATCAGAGATCTGTG 1113
Qy      280 TyrProGluGlnValLeuValSerGlyHisLeuAspSerTTPAspValGlyGlnGly 299
Db      1114 GAACCTACAGAGTATGTATTTATCTGGAGGTACCCGGAATCTCTGGATTT-----GGA 1167
Qy      300 AlaMetAspAspGlyGlyGlyAlaPheIleSerTrpIle-----AlaLeuSerLeu 316
Db      1168 GCTATTTGACCCCAACACAGTGGGGTGTCTGTTTTCAGAAATAATGCCCCGAGATTTTGAAAA 1227
Qy      317 IleLysAspLeuGlyLeuArgProLysArgThrIleuArgLeuValIleuTrpThrAlaGly 336
Db      1228 CTGATGAGTAAAGCTGAGACCTTGAAGAAATATATTTTTCAGACCTGGATGACGAA 1287
Qy      337 GluGlnGlyGlyValGlyAlaPheGlnIleTyrlGlnIleuHisLysValAsnIleSerAsn 356
Db      1288 GAAATTTGACTTCTGGGTTCACAGATGGGCTGAG-----GAGAAAT 1329
Qy      357 TyrlSerLeuValMetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThr 376

```

```

Db 1330 GTCAAAATCTCCAGGAGAGACATT---GCTTATATCAACTGGATTCTATAGAA 1386
Qy 377 GlySerGluValAlaArgAla-----11MetGluGluValMetSerLeu 391
Db 1387 GGCAATTATATCTCTCAGAGTGAAGTACTCCCTCTTTTACCAATTAGTGTATAACTG 1446
Qy 392 LeuGlnProLeu-----395
Db 1447 ACAAAAGAGATCCCGAGCCCTGATGAGTTTGAAGTAATCATCTGTATGAAGCTGG 1506
Qy 396 -----Asn1LeuGlnValLeuSerHisGlyGlu 405
Db 1507 TTGGAAGAAAGACCTTCACTGAAATTAATAATTGCTGATCAATCAATTAAGCTGGATCT 1566
Qy 406 GlyThrAsp1LeuAsnProTyr1LeuGln-----AlaGlyValProGluVala----- 420
Db 1567 GGAAGTGACTTTGAAGCTTATTTTCAAGAGACTTGAATGTCTTACGAGCAGAGCCCTTAC 1626
Qy 421 -----SerLeuLeuAspSerLeuTyr1LeuTyrPhePhePheHis1SerHisGlyAsp 438
Db 1627 ACTAAGATTAAGAAAGACAGATTAAGTACGACAGTACACAGTGTACACAAATTTATGAG 1686
Qy 439 ThrMetThrValMet-----AspPro-----LysGlnMetAsnValAla 451
Db 1687 ACATTTGAATTGGTAGAGAAATTTTATGACCCCACTTAATAAAACAATTTCTGTGGCT 1746
Qy 452 AlaAlaVal1TrrPalaValAlaSerTyrValValAlaAspMetGluGlnMetLeuPro 470
Db 1747 CAATTACGAGAGAGACAGCTGTR---TATGAGCTTGTGAT---TCTAAATCATTTCTT 1797

RESULT 10
US-10-252-157-9
; Sequence 9, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 415650.5
; US-10-252-157-9

Alignment Scores:
Pred. No.: 8,94e-10 Length: 2655
Score: 169.00 Matches: 105
Percent Similarity: 32.65% Conservative: 71
Best Local Similarity: 19.48% Mismatches: 174
Query Match: 6.99% Indels: 189
DB: 13 Gaps: 22

US-09-745-763-36 (1-472) x US-10-252-157-9 (1-2655)
Qy 3 PheLeu1IlePheAlaPhePheGlyGlyValHisLeuSerLeuCySerGlyValAla 22
Db 356 TTCTTCTCTCCGCGCTTCTCTCGGGGTGTTTAAATCTCCATGAAGACTACTAAC 415
Qy 23 IleCysLysAsnGlyIleSerLysArgThrPheGluGluIleValGluGluIleAlaSer 42
Db 416 ATTACTCCAAAGCATATATATGAAGCATTTTGGAGTGATGAAAGCTGAG----- 466
Qy 43 CysGlyAspValAlaValAlaIleIleAsnLeuAlaValTyrGlyValAlaGlnAsnArg 62

```

```

Db 467 -----AACATCAAGAGTCTTATATTAATT----- 493
Qy 63 SerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGlySerLys 82
Db 494 -----ACACAGATACCAATTTAGCGAAGACGAA 523
Qy 83 -----AsnLeuGluValAlaIleGln11MetTyrGlnAsnLeuGlnGlnAspGly 99
Db 524 CAAACTTCAAGCTTGCAAGCAATTCATCCAGTGAAGAAATTTGGCCCTGATCT 583
Qy 100 LeuGluValValHisLeuGluProVal-----Arg1LeuProHisTrr 113
Db 584 GTTGAGCTACACATTAATGATGATGCTGCTGCTCCACCAATTAAGATCACTCCCACTAC 643
Qy 114 -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
Db 644 ATCTCAATTAATTAATGAAGATGAAGAAATGGAATTTTCAACATCATTAATTGAACCA--- 700
Qy 127 IleHisLysIleAlaIleLeuGlyLeuGlySerSer1LeuGlyThrProProGluGlyIle 146
Db 701 -----CCTCCTCCAGATAT 715
Qy 147 ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnArgAlaSerGluAla 166
Db 716 GAAATGTTCGATATATGTAACACACCTTCAGTGTCTTCTCCTCAAGAAATG---CCA 772
Qy 167 ArgGlyLysIleValValTyrAsnGlnProTyr11LeuSerTyrSerArgThr 183
Db 773 GAGGCGATCTAGTG-----TATGTTAACTATGACGACACTGCAAGACTTC 817
Qy 184 -----ValGlnTyrArgThrGlnGlyValAlaGluAlaIleVal 197
Db 818 TTTAATATGGAACGGACATGAATAATGCTGCTCGGAAATTTGATATGCCAGATAT 877
Qy 198 Gly-----AlaLeuAlaSerLeu1LeuArgSerVal 207
Db 878 GGGAAAGTTTTCAGAGGAATAAGTTTAAATGCGCAGCTGCGAGGCGCAAGAGAGTC 937
Qy 208 AlaSerPheSer-----11LeuTyrSerProHisThrGlyIleGlnGluTyrGln 223
Db 938 ATTCTTACTCCAGACCTGCTGACTACTTGTCTCT---GGGAGAAAGTCTATCCA 991
Qy 224 Asp----- 224
Db 992 GATGTTGAATCTTCTCTGAGAGTGTCGTCACAGCTGGAATATCTTAATCTGAATGT 1051
Qy 224 ----- 224
Db 1052 GCAAGAGACCTCTCACACCAAGTTTACCAGCAATGATATGCTTATAGCGTGAATT 1111
Qy 225 -----GlyValProLys11LeuProThrAlaCys11LeuThrValGluAspAlaGlu 240
Db 1112 GCAAGAGCTGTGTCTTCCAGATATTCCTGTTATCATCAATGAATGATATGACAG 1171
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleVal11LeuGlnLeuLysMetGlyAla 260
Db 1172 AAGCTCTTAAG-----AAAAGGGTGGC 1195
Qy 261 LysThrTyrProAsp----- 265
Db 1196 TCAGACACCAAGATGACAGCTGAGAGAGAAAGTCTCAAGTGCCTCAATGTTGACCT 1255
Qy 265 ----- 265
Db 1256 GCGTTTACTGGAACCTTTTCTACACAAAGATGACAGATGACATCTTACCAATGA 1315
Qy 266 -----ThrAspSerPheAsnThrValAlaGlu11LeuThrGlySerLysTyrProGluGlnVal 284
Db 1316 GTGACAGAAATTTTACATATGATATGATGATCTCTCAGAGAGACAGTGAACACAGACATAT 1375
Qy 285 ValLeuValSerGlyHisLeuAspSerTrrPAspValGlyGlnGlyAlaMetAspAspGly 304

```


Db 1376 GTCAATTCGGAGTCACCGGACATCATGCTGTTT-----GGTGTATTCACCCCTCAG 1429
 Qy 305 Gtlyglalapheliesertrpglu-----Alaleuserleuilelysaspheugly 321
 Db 1430 AGTGAGACGCTGTTGTTTCATGATAATTTGTGAGACCTTTGGAACCTGAAAAAGGAAGG 1489
 Qy 322 LeuATrProLYsArGThrLeuArGLeuValLeuTrPThAlaGluGluGlnGlyVal 341
 Db 1490 TGGACACCTGAGAAACAAATTTGTTTGCAGCTGGATGCAAGAAATTTGCTTCTT 1549
 Qy 342 GlYAlaPheGlnTrYTrGlnLeuHllylsValaAnlleSerAsnTrYSerleuValMet 361
 Db 1550 GGTTCCTACTGAGTGGACAGATTTCCTTTATATATAGAAATCAAGACTCCTTCA 1609
 Qy 362 GluserApAlaGlyThrPheLeuProthrglyLeuGlnPheThrglySerGlyVala 381
 Db 1610 GAGCCTGGCGTG--GCTTATATATATGCTGACTCATCTATAGAAAGAAACATCACTCTG 1666
 Qy 382 ArgAlaIleMerGluGluValMet--SerLeuLeuGlnProLeuAsnIleThrglnVal 400
 Db 1667 AGAGTTGATTTGACACCCCTGATGTACAGCTTGTGACAC-----AACCTAACAAAGAG 1720
 Qy 401 Leu---SerHlsglyGluGlyThrasp-IleasnPheTrpIleGlnAlaGlyVal 417
 Db 1721 CTGAAAGCCCTGATGAAAGGCTTGAAGCAATCTCTTATGAAAGTTGGACTA 1775

RESULT 11

US-09-969-708-536
 ; Sequence 536, Application US/09969708
 ; Patent No. US20020102532A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Augustus, Meena
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
 ; FILE REFERENCE: Sect8
 ; CURRENT APPLICATION NUMBER: US/09/969,708
 ; CURRENT FILING DATE: 2001-10-03
 ; PRIOR APPLICATION NUMBER: US/60/237,606
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: US/60/237,608
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: US/60/237,425
 ; PRIOR FILING DATE: 2000-10-03
 ; NUMBER OF SEQ ID NOS: 658
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 536
 ; LENGTH: 2653
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-969-708-536

Alignment Scores:
 Pred. No.: 2,096-09 Length: 2653
 Score: 166.00 Matches: 105
 Percent Similarity: 32.47% Conservative: 70
 Best Local Similarity: 19.48% Mismatches: 169
 Query Match: 6.87% Indels: 195
 DB: 10 Gaps: 23

US-09-745-763-36 (1-472) x US-09-969-708-536 (1-2653)

Qy 3 PheLeuIlePheAlaPhePheGlyGlyValHlslsLeuSerLeuCySerGlyVala 22
 Db 355 TTTCTTCTCTCGGCTTCTCTTGGGGTGTATTATAAATCTCCATGAAGACTATAC 414
 Qy 23 IleCyLyAsnGlyIleSerLyArGThrPheGluGlnIleLyGluGlnIleAlaSer 42
 Db 415 ATTACTCAAGACATATATATGAACATTTTTCATGTAATGAAGAGTGAG-- 465
 Qy 43 CysGlyAspValAlaValAlaIleleAsnLeuAlaValTrGlyLyAlaGlnAsnArg 62
 Db 466 -----AACATCAAGAGTTCTTATATATTT----- 492

Qy 63 SerTrGluArGLeuAlaLeuLeuValAspThrValGlyProArGLeuSerGlySerLy 82
 Db 493 -----ACAGAGTACCACTTTAGCAGAAACGAA 522
 Qy 83 -----AsnLeuGlyValAlaIleGlnIleMetYrGlnAsnLeuGlnAspGly 99
 Db 523 CAAACTTTCAGCTTGGAAAGCAAAATTCATCCAGTGAAGAAATTTGGCTGGAATTC 582
 Qy 100 LeuGlyLyValHlslsLeuGlnProVal-----ArgIleProHlstrp 113
 Db 583 GTTAGCTAGCACATTTATGATGATGCTGTGTCTTCAACCAATTAAGACTCATCCAACTAC 642
 Qy 114 -----GluArGlyGluGlu-----SerAlaValMetLeuGlnProArg 126
 Db 643 ATTCATATATATATGAGATGAGAAATGAGATTTTCAACATCATATTATGAAACA-- 699
 Qy 127 IleHlslsIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGlnGlyIle 146
 Db 700 -----CCTCCTCCAGATAT 714
 Qy 147 ThrAlaGluValLeuValAlaThrSerPheAspGluLeuGlnArGAlaSerGluAla 166
 Db 715 GAAATGTTTCGATATATGTAACACCTTTCAGTCTTCTCTCAAGAGATG--CCA 771
 Qy 167 ArgGlyLyIleValValTrAsnGlnProTrYrIleAsnTrYrSerArGThr----- 183
 Db 772 GAGGCGCATCTAGT-----TATCTTACTATGACCAACTCAAGACTTC 816
 Qy 184 -----ValGlnTrArGThrglnGlyAlaValGluAlaAlaLyVal 197
 Db 817 TTTAAATTTGAACCGGACATGAATAATCAATGCTCTGGGAAATTTGTAATTCACAGATAT 876
 Qy 198 Gly-----AlaLeuAlaSerLeuIleArGSerVal 207
 Db 877 GGGAAAGTTTCAGAGAAATTAAGCTTAAATGCCAGCTGAGGAGGCCAAAGAGATC 936
 Qy 208 AlaSerPheSer-----IleYrSerProHlstrhrglyIleGlnGlyTrGln 223
 Db 937 ATTCTACTACCGACCGCTGCTGACTTGTGCTCT-----GGGGAAGTCTCATCA 990
 Qy 224 Asp----- 224
 Db 991 GATGTTGAATCTTCTGAGAGTGTGTCCAGCGTGAATAATCTTAATCTGAATGCT 1050
 Qy 224 ----- 224
 Db 1051 GCAGAGACCTCTCACACAGGTTACCCAGCAATGATATGCTTATAGCGGTGAATT 1110
 Qy 225 -----GlyValProLyIleProThrAlaCyIleThrValGluAspAlaGlu 240
 Db 1111 GCAGAGCTGTGTCTTCCAGATATCTGTTTCATCATCAATGATGATCTATGATGACAG 1170
 Qy 241 MetWetSerArGMetAlaSerHlsglyIleLyIleValIleGlnLeuLyMetGlyAla 260
 Db 1171 AAGCTCTTAGA-----AAAATGGGTGGC 1194
 Qy 261 LysThrYrProAsp----- 265
 Db 1195 TCAGCACCCACCATGACCTGACGTGAGAGAGATCTCAAGTCCCTTAATGTTGACCT 1254
 Qy 265 ----- 265
 Db 1255 GGCCTTACTGAAACTTTTCTACCAAAAAAGTCAAGATGACATCCACTTACCAATGAA 1314
 Qy 266 ---ThrasPheAsnThrValAlaGluIleThrglySerLyTrProGluGlnVal 284
 Db 1315 GTGACAGAAATTTACATGATGATGATGATCTCTCAGAGACAGAGCAACAGATAT 1374
 Qy 285 ValLeuValSerGlyHlslsLeuAspSerTrpAspValGlyGlnGlyAlaMetAspGly 304
 Db 1375 GTCAATTTGGAGGTTCACCGGACTCATGGGTGTT-----GGTGAATTCACCTTCAG 1428
 Qy 305 Gtlyglalapheliesertrpglu-----Alaleuserleuilelysaspheugly 321

Db 1429 AGTGGAGAGAGCTGTTCATGAAATTCGTGAGAGCTTTGGACACCTGAAAAGGAGG 1488
 Qy 322 LeuArgProLysArgThrLeuArgLeuValLeuThrPheArgLeuGluGlnGlyVal 341
 Db 1489 TGGAGACCTAGAACAAATTTGTTGCAAGCTGGAGTGAAGAAATTTGGCTCTCTT 1548
 Qy 342 GylAlaPheGlnTyrTyrGlnLeuHisLeuValAsnLeuSerAsnTyrSerLeuValMet 361
 Db 1549 GGTTCCTACTAGCTGGCAGAG-----GAGAAATTCAGACCTCTTCAA 1590
 Qy 362 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluValAla 381
 Db 1591 GAGCGTGGCGT---GCTTATATTATGCTACTCATCTATAGAAAGAACTACACTCTG 1647
 Qy 382 ArgAlaIleMetGluGluValMet---SerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 Db 1648 AGAGTTGATTGTACACCGCTGATGTACAGCTTGACAC-----ACCTACAGAAAGAG 1701
 Qy 401 Leu---SerHisGlyGluGlyThrAsp-IleAsnPheTrpIleGlnAlaGlyVal 417
 Db 1702 CTGAAAAGCCTGATGAAAGCCTTGAGAGCAATCTCTTATGAAAGTTGAGACTA 1756
 RESULT 12
 US-09-873-319-451
 ; Sequence 451, Application US/09873319A
 ; Publication No. US20030134324A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Munger, William E.
 ; APPLICANT: Kulkarni, Prakash
 ; APPLICANT: Getzenberg, Robert H.
 ; APPLICANT: Waga, Iwao
 ; APPLICANT: Yamamoto, Jun
 ; TITLE OR INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
 ; FILE REFERENCE: 44921-5029-US
 ; CURRENT APPLICATION NUMBER: US/09/873,319A
 ; EARLIER FILING DATE: 2001-06-05
 ; EARLIER APPLICATION NUMBER: US 60/223,323
 ; EARLIER FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 755
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 451
 ; LENGTH: 2653
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M99487
 ; US-09-873-319-451
 Alignment Scores:
 Pred. No.: 2,09e-09 Length: 2653
 Score: 166.00 Matches: 105
 Percent Similarity: 32.47% Conservative: 70
 Best Local Similarity: 19.48% Mismatches: 169
 Query Match: 6.87% Indels: 195
 DB: 13 Gaps: 23
 US-09-745-763-36 (1-472) x US-09-873-319-451 (1-2653)
 Qy 3 PheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGlyValAla 22
 Db 355 TTCTTTCTCCCTCGGCTTCCTCTCGGTGTTAAATAATCTCCAAAGAAAGCTACTAC 414
 Qy 23 IleCysLysAsnGlyIleSerLysArgThrPheGluGlnIleLysGluIleAlaSer 42
 Db 415 ATTACTCCAAAGCAATATATGAAAGCATTTTGGATGATGAAAGCTGAG----- 465
 Qy 43 CysGlyAspValAlaLysAlaIleIleLeuLeuAlaValIleGlyLysAlaGlnAsnArg 62
 Db 466 -----AACATCAAGAGTTCTTATATATTT----- 492
 Qy 63 SerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGlySerLys 82

Db 493 -----ACACAGATACCACTTAGCAGAAACGAA 522
 Qy 83 -----AsnLeuGlyValAlaIleGlnIleMetLysTyrGlnAsnLeuGlnAsnGly 99
 Db 523 CAAAACTTCAGCTTGAAAGCAAAATTCATCCAGGAAAGAAATTTGGCTGGATTC 582
 Qy 100 LeuGluLysValHisLeuGluProVal-----ArgIleProHisTrp 113
 Db 583 GTTAGCTAGACACTTATGATGTCCTTGTGCTTCAACCAATAGATCATCCCACTAC 642
 Qy 114 -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
 Db 643 ATTCATTAATTAATGAAGATGAAGATGATTTTCAACATCATTAATTGAACCA--- 699
 Qy 127 IleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146
 Db 700 -----CCTCTCCAGAGATAT 714
 Qy 147 ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnArgArgAlaSerGluAla 166
 Db 715 GAAATGTTCCGATATGTACACCTTCAGTCCTTCTCTCTCAAGAAATG---CCA 771
 Qy 167 ArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyrSerArgThr----- 183
 Db 772 GAGGGCATCTAGTG-----TATGTTAACTATGACCACTGAAGACTTC 816
 Qy 184 -----ValGlnTyrArgThrGlnGlyValAlaGluAlaIleVal 197
 Db 817 TTTAAATGGACGGACATGAATCAATGCTCTCGGAAATGTAATGATCCAGATAT 876
 Qy 198 Gly-----AlaLeuAlaSerLeuIleArgSerVal 207
 Db 877 GGGAAAGTTTTCAGAGAAATTAAGCTTAAATGCCCCAGCTGGAGGGCCAAAGAGTTC 936
 Qy 208 AlaSerPheSer-----IleTyrSerProHisThrGlyIleGlnGluTyrGln 223
 Db 937 ATTCTACTCCACACCTGCGACTACTTGTCTCT-----GGGGAAGTCTTATCCA 990
 Qy 224 Asp----- 224
 Db 991 GATGTTGAATCTTCTCGAGAGTGTCGTCACAGCTGAGAAATATCTTAATCTGAATGCT 1050
 Qy 224 ----- 224
 Db 1051 GCAGAGACCTCTCACACCAAGTTACCCAGCAATGATATGCTTATAGCGGTGAATT 1110
 Qy 225 -----GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 Db 1111 GCAGAGCTGTGGTCTTCCAGATTCCTGTTCAATCCATGATGATGATGACAG 1170
 Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 Db 1171 AAGCTCTTAGA-----AAAATGGGTGGC 1194
 Qy 261 LysThrTyrProAsp----- 265
 Db 1195 TCAGCACCAACCAATGACAGCTGAGAGGAAGTCTCAAGTCCCTTAATGTTGACCT 1254
 Qy 265 ----- 265
 Db 1255 GCGTTTACTGGAACCTTTCTACACAAAGTCAAGATGACATCCACTACCAATGAA 1314
 Qy 266 ---ThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyrProGluGlnVal 284
 Db 1315 GTGACAAAGAAATTTACATATGATGATGATCTCTGAGAGACAGTGAACAGACAGATAT 1374
 Qy 285 ValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGly 304
 Db 1375 GTCAATTCGGAGGTCAACCGGACTCAAGGCTGTT-----GCTGGTATTTGACCTCAG 1428
 Qy 305 GlyGlyAlaPheIleSerTrpGlu-----AlaLeuSerLeuIleLysAspLeuGly 321

Db 1429 AGTGAGACGCTGTTGTCATGAAATTTGAGACCTTTGGAACCTGAAAAAGAGAGG 1488
Qy 322 LeuArgProLysArgThrLeuArgLeuValLeuTPrMAlaGluGluGlnGlyVal 341
Db 1489 TGGAGACCTGAGAGAACTTTGTTTGGACCTGGATGACAGAAATTTGGTCTTCTT 1548
Qy 342 G1yAlaPheGlnTyrTrgInLeuHisLysValAsnIleSerAsnTyrSerLeuValMet 361
Db 1549 GGTCTACTAGAGTGGCAGAG-----GAGAATTCAGAGACTCTTCA 1590
Qy 352 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyVal 381
Db 1591 GAGCGTCGCGTG--GCTTATATTAATGCTGACTCATCTATAGAGAAACTACCTCTG 1647
Qy 382 ArgAlaIleMetGluGluValMet--SerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db 1648 AGAGTTGATGTACACCCGCTGATGTACAGCTTGATCAC-----AACCTACAAAAGAG 1701
Qy 401 Leu--SerHisGlyGluGlyThrAsp-IleAsnPheTrpIleGlnAlaGlyVal 417
Db 1702 CTGAAAGCCCTGATGAAAGGCTTTGAGGCAATCTTTATGAAAGTTGACATA 1756

RESULT 13
US-09-960-706-716
; Sequence 716, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 716
; LENGTH: 2653
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M99487
US-09-960-706-716

Alignment Scores:
Pred. No.: 2,096-09 Length: 2653
Score: 166.00 Matches: 105
Percent Similarity: 32.47% Conservative: 70
Best Local Similarity: 19.48% Mismatches: 169
Query Match: 6.87% Indels: 195
DB: 13 Gaps: 23

US-09-745-763-36 (1-472) x US-09-960-706-716 (1-2653)

Qy 3 PheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGlyValAla 22
Db 355 TTCCTTCTCTCGGCTTCTCTTTCGGGTGTTTAAATAATCTCCAAATGAACCTACTAAC 414
Qy 23 IleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIleAlaSer 42
Db 415 ATTACTCCAAACCAATTAATATGAAAGCATTTTGGATGATTAAGAGCTGAG----- 465
Qy 43 CysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGlnAsnArg 62
Db 466 -----AACATCAAGAAAGTTTATATATATT----- 492
Qy 63 SerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGlySerLys 82
Db 493 -----ACACAGATACCAACATTTAGCAGAAACAGAA 522

Qy 83 -----AsnLeuGlyValAlaIleGlnIleMetTyrTrgInLeuLeuGlnAlaAspGly 99
Db 523 CAAACTTTGCTGCAAGAAACAAATTCATCCAGATGAAAGAAATTTGGCTGGATCT 582
Qy 100 LeuGlyLysValHisLeuGluProVal-----ArgIleProHisTrp 113
Db 583 GTTAGCTAGCACAATTATGATGATGCTGCTGCTTCCACCAATAAGACTATCCACATAC 642
Qy 114 -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
Db 643 ATTCATTAATTAATGAAATGATGAAATGAGATTTTCAACATCATATTGAAACCA--- 699
Qy 127 IleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146
Db 700 -----CTCTCCAGAGATAT 714
Qy 147 ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnAlaGlyAlaSerGluAla 166
Db 715 GAAATGTTTCGATATGTATGACACCTTCAGTCTTCTCTCCACAGAAAG--CCA 771
Qy 167 ArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyrSerArgThr----- 183
Db 772 GAGGCGATCTAGT-----TATGTATCTATGACACGAACTGAAAGACTTC 816
Qy 184 -----ValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysVal 197
Db 817 TTTAAATTTGAAACGAGCAATGAAATCAATGCTCTGGAAATATGTAATGCCAGATAT 876
Qy 198 Gly-----AlaLeuAlaSerLeuIleArgSerVal 207
Db 877 GGGAAAGTTTTCAGAGAAATTAAGTTAAATATCCAGCTGCGAGGGGCCAAAGAGATC 936
Qy 208 AlaSerPheSer-----IleTyrSerProHisThrGlyIleGlnGluTyrGln 223
Db 937 ATTCTTACTCCGACCTGCTGACTACTTGTCTCT-----GGGTGAAGTCTTATCCA 990
Qy 224 Asp----- 224
Db 991 GATGGTTGAATCTTCTCGAGAGTGTGTCCAGCGTGAATATCTTAATCTGAATGAT 1050
Qy 224 ----- 224
Db 1051 GCAGAGACCTCTACACACGATTACCCAGCAATGATATGCTTATAGCGTGAAAT 1110
Qy 225 -----GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 1111 GCAGAGCGCTGTGATCTTCCAGATATCTCTTATCTCAATGGATATGATATGATGACAG 1170
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 1171 AAGCTCTAGAA-----AAATGGGTGGC 1194
Qy 261 LysThrTyrProAsp----- 265
Db 1195 TCACACACACAGATGACAGCTGAGAGAGAGTCTCAAGTCCCTCAATGTTGACCT 1254
Qy 265 ----- 265
Db 1255 GCGTTTACTGAAACTTTTCTACCAAAAAGTCAAGATGCACATCTACCAATGA 1314
Qy 266 ---ThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyrProGluGlnVal 284
Db 1315 GTGACAGAAATTTACATGATATGATATGATCTGACAGAGAGAGAGCAACGACAGATAT 1374
Qy 285 ValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGly 304
Db 1375 GTCAATCTGGAGGCTCACCGGACTCATGGGTCTT-----GGTGGTATGACCTCAG 1428
Qy 305 GlyGlyAlaPheIleSerTrpGlu-----AlaLeuSerLeuIleLysAspLeuGly 321
Db 1429 AGTGAGACGCTGTTGTCATGAAATTTGAGACCTTTGGAACCTGAAAAAGAGAGG 1488
Qy 322 LeuArgProLysArgThrLeuArgLeuValLeuTrpMAlaGluGlnGlyVal 341

```

Db      1489 TGGAGACCTAGACAAACATTTTGTTCAGAGCTGGAGTCAGAGAAATTTGGCTTCTT 1548
Oy      342 G1yAlahegInTyTgInLeuHieLyValAsnIleSerAntyTserLeuValMet 361
Db      1549 GGTCTACTAGTGCGACAG------GAGAAATTCAGAGCTCTCCAA 1590
Oy      362 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyValAla 381
Db      1591 GAGCGTGCGCTG---GCTATATTATATGCTGATCATCATATAGAGAAACTACACTCTG 1647
Oy      382 ArgAlaIleMetGluGluValMet---SerLeuGlnProLeuAsnIleThrGluVal 400
Db      1648 AGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACAC-----AACCTAACAAAGAG 1701
Oy      401 Leu---SerHisGlyGluGlyThrAsp-IleAsnPheThrIleGlnAlaGlyVal 417
Db      1702 CTGAAAGCCTGATGAAAGCTTTGAAAGCAATCTCTTATGAAAGTTGAGACTA 1756

```

RESULT 14

```

US-10-210-120-5
; Sequence 5, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
; APPLICANT: Chinnaiyan, Arul M.
; APPLICANT: Rubin, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UM-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/334,468
; PRIOR FILING DATE: 2001-11-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 2653
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-210-120-5

```

Alignment Scores:

```

Pred. No.:      2,09e-09      Length:      2653
Score:          166.00      Matches:      105
Percent Similarity: 32.47%      Conservative: 70
Best Local Similarity: 19.48%      Mismatches: 169
Query Match:      6.87%      Indels:      195
                        Gaps:      23

```

US-09-745-763-36 (1-472) x US-10-210-120-5 (1-2653)

```

Oy      3 PheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGlyValysAla 22
Db      355 TTCTTTCTCCTCGGCTTCTCTGCGGTGTTATAAATCTCCAAATGAGACACTAC 414
Oy      23 IleCySerHisGlyIleSerIleArgThrPheGluGluIleGlyGluIleAlaSer 42
Db      415 ATTACTCCAAAGCAATATATGAAAGCATTTTGGATGTAATGAAAGCTGAG- 465
Oy      43 CysGlyAspValAlaIleValAlaIleLeuAsnLeuAlaValTyGlyValAlaGlnAsnArg 62
Db      466 ---AACATCAAGAGCTCTTATATATTTT----- 492
Oy      63 SerTyGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGlySerIys 82
Db      493 ---ACACGATACCACTTTAGCAGAAACAGAA 522
Oy      83 ---AsnLeuGluValAlaIleGlnIleMetTyGlnAsnLeuGlnGlnArgGly 99
Db      523 CAAAACTTTCAGCTTGCAAGCAAAATTCATCCAGTGAGAAAGAAATTTGGCTGATTTCT 582

```

```

Oy      100 LeuGlyValHisLeuGluProVal-----ArgIleProHisTyr 113
Db      583 GTTGAGCTAGACATTAATGATGTCCTGTGCTTACCCAAATAAGACTATCCCACTAC 642
Oy      114 ---GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
Db      643 ATCTCAATTAATTAAGTAAGTAAGTAAGTAAGTAATTTTCACATCATTAATTTGAAACA 699
Oy      127 IleHisValIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146
Db      700 ---CTCTCCAGGATAT 714
Oy      147 ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnArgArgAlaSerGluAla 166
Db      715 GAAATGTTCCGATATGTATGACCACTTACCACTTACGCTTCTCTCTCCTCAAGGAATG---CCA 771
Oy      167 ArgGlyLysIleValValTyraGlnProTyrlleAsnTySerArgThr----- 183
Db      772 GAGGCGCATCTAGT---TATGTTAACTATGACGACGACGAAAGACTTC 816
Oy      184 ---ValGlnTyraGThrGlnGlyAlaValGluAlaIleVal 197
Db      817 TTTAATTTGAGACGGACATGAAATCATATGCTCTGGAAATTTGATATGCGAGATAT 876
Oy      198 Gly-----AlaLeuAlaSerLeuIleArgSerVal 207
Db      877 GGGAAATTTTCAGAGAAATTAAGTTAAATATGCCAGCTGCGCGCAAGAGATC 936
Oy      208 AlaSerPheSer-----IleTySerProHisThrGlyIleGlnIleTyGln 223
Db      937 ATTCTACTCCACCCCTGCTGACTACTTGTCTCT---GGGGAAGTCTTATCCA 990
Oy      224 Asp----- 224
Db      991 GATGTTGGAATCTTCTGAGAGTGCTGCCAGCGTGAATATCTAAATTCGAATGCT 1050
Oy      224 --- 224
Db      1051 GCAGAGACCTCTACACACAGGTATCCAGCAATGATATGCTTATAGCGGTGAAT 1110
Oy      225 ---GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      1111 GCAGAGCTGTGCTTCCAGATATCTGTTATCATCATGATGATGATGATGACAG 1170
Oy      241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db      1171 AAGCTCTAGAA-----AAAATGGGTGGC 1194
Oy      261 LysThrTyProAsp----- 265
Db      1195 TCAGCACCAACAGTATGACAGTGAAGAGAAAGTCTCAAAGTCCCTTCAATGTGGACCT 1254
Oy      265 --- 265
Db      1255 GGCTTACTGAAACTTTTTCACACAAAAGTCAAGATGACATCCACTACCAATGAA 1314
Oy      266 ---ThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyProGluGlnVal 284
Db      1315 GTGACAAAGAAATTAACAATGATAGTACTCTCAGAGACAGTGAACAGACAGATAT 1374
Oy      285 ValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAlaMetAspArgGly 304
Db      1375 GTCATTTTGGGAGGTACCGGAGCTCATGGGTGTTT-----GGTGGTATTTGACCTTCAG 1428
Oy      305 GlyGlyAlaPheIleSerTyrGlu-----AlaLeuSerLeuIleLysAspLeuGly 321
Db      1429 AGTGAGACAGCTGTGTTTCATGAATGTGAGAGGCTTTGAAACACTGAAAGAAAGAGGG 1488
Oy      322 LeuArgProLysArgThrIleuArgLeuValLeuTyThrAlaGluGluGlnGlyVal 341
Db      1489 TGGAGACCTTAGAGAAATTTTGTGCAAGCTGGGATCCAGAAAGAAATTTGGCTTCTT 1548
Oy      342 GlyAlaPheGlnTyTgInLeuHieLyValAsnIleSerAntyTserLeuValMet 361

```

```

Db      1549 GGTCTACTAGTGGGACAG-----GAGAAATTCAGAACTCTCTCA 1590
Qy      362 GUSERAEPALAGLYThrPheLeuProThrGlyLeuGlnPheThrGlySerLeuIysAla 381
Db      1591 GAGCGTGGCGTG---GCTTATATTATAGCTGACTCATCTATGAAAGAACTACACTCTG 1647
Qy      382 ARGAlaIleMetGluGluValMet---SerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1648 AGAGTTGATTTGACACCGCTGATGTCACCTTGTCACAC-----AACCTAACAAAGAG 1701
Qy      401 Leu---SerHisGlyGluGlyThrAsp-IleAsnPheTrpIleGlnAlaGlyVal 417
Db      1702 CTGAAAGCCCTGATGAAAGCTTTGAAAGCAATCTCTTATGAAAGTTGACTA 1756

RESULT 15
US-10-341-434-70
/ Sequence 70, Application US/10341434
/ Publication No. US20030215835A1
/ GENERAL INFORMATION:
/ APPLICANT: Origene Technologies
/ TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
/ FILE REFERENCE: 9U 204 205 R1
/ CURRENT APPLICATION NUMBER: US/10/341,434
/ PRIOR FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: US 60/348,164
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: US 60/348,119
/ NUMBER OF SEQ ID NOS: 238
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 70
/ LENGTH: 2653
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (262..(2511)
/ OTHER INFORMATION:
US-10-341-434-70

Alignment Scores:
Pred. No.: 2,09e-09 Length: 2653
Score: 166.00 Matches: 105
Percent Similarity: 32.47% Conservative: 70
Best Local Similarity: 19.48% Mismatches: 169
Query Match: 6.87% Indels: 195
DB: 13 Gaps: 23

US-09-745-763-36 (1-472) x US-10-341-434-70 (1-2653)
Qy      3 PheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGlyIysAla 22
Db      355 TTCCTTCTCCCTCGGCTTCTCTCGGCTGTTTATTAATAATCCTCCAAATGAAGCTACTAC 414
Qy      23 IleCyIysAsnGlyIleSerIysArGThrPheGluGluIleIysGluGluIleAlaSer 42
Db      415 ATTACTCCAAAGCAATATATGAAAGCATTTTGGATGATGAAAGCTGAG-----465
Qy      43 CySGIysPValAlaIleValIleIleAsnLeuAlaValIyGlyIysAlaGlnAsnArg 62
Db      466 -----AACATCAAGAGTTCTTATATAATTT-----492
Qy      63 SerIyGluIyGluLeuAlaLeuLeuValAspThrValGlyProIyGLeuSerGlySerIys 82
Db      493 -----ACACAGATACCCACATTTGACGAGAACAA 522
Qy      83 -----AsnLeuGluIysAlaIleGlnIleMetCyTgIAsnLeuGlnGlnAspGly 99
Db      523 CAAACCTTCAAGCTTGCAGAAAGCAATTCATCCAGTGAAGAAATTTGGCTGATTCCT 582
Qy      100 LeuGluIysValHisLeuGluProVal-----ArgIleProHisTrp 113

```

```

Db      583 GTTAGCTAGACATTAATGATGCTGTGTCTTACCCAAATTAAGTATCCCAACTAC 642
Qy      114 -----GluArgIyGluGlu-----SerAlaValMetLeuGluProArg 126
Db      643 ATTCATATTAATTAATGAAGATGAAATGATTTTCAACATCATTAATTTGAACCA---699
Qy      127 IleHisIysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146
Db      700 -----CCTCCCTCCAGGATAT 714
Qy      147 ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnArgAlaSerGluAla 166
Db      715 GAAATGTTTGGATATGTATGACCACTTCAGTCTCTCTCCCAAGAAATG---CCA 771
Qy      167 ArgGlyIysIleValIyTrAsnGlnProIyTrIleAsnIySerArgThr-----183
Db      772 GAGGCGATCTAATG---TATGTTAACTTAATGACAGAACTGAAGACTTC 816
Qy      184 -----ValGlnIyTrArgThrGlnGlyAlaValAlaIyVal 197
Db      817 TTTAAATTTGAACGGACATGAATAATCATTCCTGGAAATTTGAATTTGCCAGATAT 876
Qy      198 Gly-----AlaLeuAlaSerLeuIleArgSerVal 207
Db      877 GGGAAAGTTTTCAGAGAAATTAAGTTAAATCCCAAGCTGCGAGGGGCCAAAGAGATC 936
Qy      208 AlaSerPheSer-----IleTrSerProHisTrpIleGlnIyGlnIy 223
Db      937 ATTCTTACTCCGACCTGCTGACTGACTTGTGCTCT---GGGATGAAGTCTTATCCA 990
Qy      224 Asp-----224
Db      991 GATGGTTGAATCTTCTCGAGAGTGCTGCCAGGTGGAATATCTTAATCTGAATGT 1050
Qy      224 -----224
Db      1051 GCAGAGACCTCTCACACAGTTACCAAGTAAATGATGCTTATAGCGTGAATTT 1110
Qy      225 -----GlyValProIyIleProThrAlaCySileThrValGluAspAlaGlu 240
Db      1111 GCAGAGCTGTGGTCTTCCTCAAGTATTCCTGTTATCCAAATGATATCTATGATGCAAG 1170
Qy      241 MetMetSerArgMetAlaSerHisGlyIleIysIleValIleGlnLeuIysMetGlyAla 260
Db      1171 AAGCTCTAGAA-----AAATGGGTGCC 1194
Qy      261 LysThrIyTrProAsp-----265
Db      1195 TCAGCACCAACAGATACAGCTGAGAGAGAACTCAAGTCCCTCAATGTTGACCT 1254
Qy      265 -----265
Db      1255 GGGTTTATCTGAAACTTTTCTACACAAAGTCAAGATGCACATCTACCAATGAA 1314
Qy      266 ---ThraSerPheAsnThrValAlaGluIleThrGlySerIySlyTrProGluGlnVal 284
Db      1315 CTGACAAAGATTTTCAATTCATTCATAGTACCTCAGAGAGACAGTGAACCAAGATAT 1374
Qy      285 ValLeuValSerGlyIysLeuAspSerTrpAspValIyGlnGlyAlaMetAspAspGly 304
Db      1375 GTCAATTCGGAGGCTCACCGGACTCATGGGTGTT-----GGTGTATTTGACCTCAG 1428
Qy      305 GlyIyValaPheIleSerTrpGlu-----AlaLeuSerLeuIleIysAspLeuGly 321
Db      1429 AGTGAAGCAGCTGTGTTTCATGAATTTGAGAGGCTTTGGAACATGAAAGAAAGAGG 1488
Qy      322 LeuArgProIysAspGThrLeuArgLeuValLeuTrpHisAlaGluGlnGlnIyGlyVal 341
Db      1489 TGGAGACCTTAAGAAACAAATTTTGTTCAGAGCTGGATGCAAGAAATTTGGCTTCTT 1548
Qy      342 GlyAlaPheGlnIyTrGlnIleuHisIysValAsnIleSerAsnIyTrSerLeuValMet 361
Db      1549 GGTCTACTAGTGGGACAG-----GAGAAATTCAGAACTCTCTCA 1590

```

```
Qy      362 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLysAla 381
          |||
          ::::
Db      1591 GAGCGTGGCCG---GCTTATATTAAAGCTGACTCATCTATAGAAAGAACTACACTCTG 1647
          |||
Qy      382 ArgAlaIleMetGluGluValMet--SerLeuLeuGlnProLeuAsnIleThrGluVal 400
          |||
          ::::
Db      1648 AGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACAC-----AACCTAACAAAGAG 1701
          |||
Qy      401 Leu--SerHisGlyGluGlyThrAsp-IleAsnPheTrpIleGlnAlaGlyVal 417
          |||
          ::::
Db      1702 CTGAAAAGCCCTGATGAGGCTTTGAAGCAAAATCTTATGAAAAGTTGACTA 1756
          |||
```

Search completed: December 23, 2003, 01:53:39
Job time : 435 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:53:06 / Search time 35 Seconds
(without alignments)
73.729 Million cell updates/sec

Title: US-09-745-763-36_COPY_231_240
Perfect score: 10
Sequence: 1 TACTIVEDAE 10

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 0
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	100.0	433	11 070216	Q70216 mus musculu
2	10	100.0	470	11 09WVJ3	Q9WVJ3 mus musculu
3	10	100.0	472	4 Q8NBZ1	Q8NBZ1 homo sapien
4	10	100.0	472	4 Q9Y646	Q9Y646 homo sapien
5	10	100.0	541	4 Q9Y5X6	Q9Y5X6 homo sapien
6	7	70.0	576	11 Q8BGT9	Q8BGT9 mus musculu
7	6	60.0	157	10 Q9SWJ1	Q9SWJ1 arabidopsis
8	6	60.0	167	2 Q9LJ90	Q9LJ90 vibrio chol
9	6	60.0	171	16 Q9KMC6	Q9KMC6 vibrio chol
10	6	60.0	185	2 Q8KVK6	Q8KVK6 uncultured
11	6	60.0	186	2 Q8KVC0	Q8KVC0 uncultured
12	6	60.0	186	2 Q8KVG2	Q8KVG2 uncultured
13	6	60.0	186	2 Q8KHQ8	Q8KHQ8 uncultured
14	6	60.0	186	13 Q9PSY2	Q9PSY2 gallus gall
15	6	60.0	187	16 Q9ABZ8	Q9ABZ8 caulobacter
16	6	60.0	188	11 Q8K4X9	Q8K4X9 mus musculu

17	6	60.0	197	16 Q9KM80	Q9KM80 vibrio chol
18	6	60.0	205	11 Q9LXU1	Q9LXU1 rattus norv
19	6	60.0	224	4 Q9UPF5	Q9UPF5 homo sapien
20	6	60.0	230	4 Q9S553	Q9S553 homo sapien
21	6	60.0	246	4 Q9S552	Q9S552 homo sapien
22	6	60.0	260	2 Q52442	Q52442 pseudomonas
23	6	60.0	261	2 Q9Z158	Q9Z158 pseudomonas
24	6	60.0	271	11 Q9R2A8	Q9R2A8 mus musculu
25	6	60.0	281	16 Q8PJZ8	Q8PJZ8 xanthomonas
26	6	60.0	293	4 Q9E6J3	Q9E6J3 homo sapien
27	6	60.0	299	11 Q9CW34	Q9CW34 mus musculu
28	6	60.0	309	17 Q8PYR5	Q8PYR5 methanosarc
29	6	60.0	315	4 Q9P0X8	Q9P0X8 homo sapien
30	6	60.0	319	4 Q96PU6	Q96PU6 homo sapien
31	6	60.0	319	4 Q96PU7	Q96PU7 homo sapien
32	6	60.0	319	11 Q9QUH4	Q9QUH4 mus musculu
33	6	60.0	321	4 Q9P0Y1	Q9P0Y1 homo sapien
34	6	60.0	323	4 Q96KA3	Q96KA3 homo sapien
35	6	60.0	325	4 Q969L9	Q969L9 homo sapien
36	6	60.0	325	11 Q61110	Q61110 mus musculu
37	6	60.0	328	16 Q8G3H1	Q8G3H1 bifidobacte
38	6	60.0	330	2 Q9KM44	Q9KM44 wolbachia s
39	6	60.0	330	2 Q9KM37	Q9KM37 wolbachia s
40	6	60.0	330	11 Q8K4Y0	Q8K4Y0 mus musculu
41	6	60.0	332	2 Q8RPM0	Q8RPM0 ehrlichia c
42	6	60.0	332	13 Q9YH19	Q9YH19 gallus gall
43	6	60.0	333	16 Q910X1	Q910X1 pseudomonas
44	6	60.0	334	2 Q8GLV8	Q8GLV8 ticketstia
45	6	60.0	334	16 Q921N0	Q921N0 ticketstia

ALIGNMENTS

RESULT 1

ID	070216	PRELIMINARY;	PRT;	433 AA.
AC	070216;			
DT	01-AUG-1998 (TREMBLERL. 07, Created)			
DT	01-AUG-1998 (TREMBLERL. 07, Last sequence update)			
DT	01-MAR-2003 (TREMBLERL. 23, Last annotation update)			
DB	Hematopoietic lineage switch 2.			
GN	PGCP OR HLS2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
NC	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Williams J.H., Chan C.-Y., Klinken S.P.,			
RT	"Hematopoietic lineage Switch 2 (HLS2), a novel mRNA species induced			
RT	during an erythroid to myeloid lineage switch."			
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF009513; AAC17945.1; -			
DR	MEROPS; M28.014; -			
DR	MCD; MGI:1869205; Pgcp.			
DR	InterPro; IPR001464; Annexin.			
DR	InterPro; IPR003137; PA.			
DR	PROSITE; PS00223; ANNEXIN; 1.			
DR	PROSITE; PS50840; PA; 1.			
SQ	SEQUENCE 433 AA; 47858 MW; 89C9B93712004669 CRC64;			

Query Match 100.0%; Score 10; DB 11; Length 433;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACTIVEDAE 10
|||
Db 229 TACTIVEDAE 238

RESULT 2
Q9WVJ3

```

ID Q9WVJ3 PRELIMINARY; PRT; 470 AA.
AC Q9WVJ3:
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Aminoacyltransferase (similar to plasma glutamate carboxypeptidase).
GN PCGP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Liu C.H., Lin B.Y., Chang L.Y.;
RT "Cloning of the mouse aminopeptidase gene."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strauberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body, and Mesonephros;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
DR EMBL; AF107835; AAD43215.1; -
DR EMBL; BC037067; AAH37067.1; -
DR EMBL; AK032972; BAC28105.1; -
DR EMBL; AK075686; BAC35891.1; -
DR MEROPS; M28.014; -
DR MGD; MGI:1889205; Pcp.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR003137; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS50840; PA; 1.
DR Carboxypeptidase.
SQ SEQUENCE 470 AA; 51813 MW; 0F3490681691866A CRC64;

Query Match 100.0%; Score 10; DB 11; Length 470;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIVEDAE 10
DB 229 TACTIVEDAE 238

RESULT 3
Q9NBZ1 PRELIMINARY; PRT; 472 AA.
ID Q9NBZ1:
AC Q9NBZ1:
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ90651.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isegai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hiro Y., Saito K.,
RA Yamamoto Y., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Maehiro Y., Ota T., Okano K., Yoshikawa Y., Aotoku S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;

```

```

RT "NEBO human cDNA sequencing project."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075132; BAC11423.1; -
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 51784 MW; 83F2596F5D5C2F3 CRC64;

Query Match 100.0%; Score 10; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIVEDAE 10
DB 231 TACTIVEDAE 240

RESULT 4
Q9Y646 PRELIMINARY; PRT; 472 AA.
ID Q9Y646:
AC Q9Y646:
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Aminoacyltransferase (Plasma glutamate carboxypeptidase).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Liu C.H., Lin B.Y., Chang L.Y.;
RT "Cloning of the human aminopeptidase gene."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strauberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF107834; AAD43214.1; -
DR EMBL; BC020689; AAD20689.1; -
DR MEROPS; M28.014; -
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
DR Carboxypeptidase.
SQ SEQUENCE 472 AA; 51887 MW; BE6CBD2149E042BF CRC64;

Query Match 100.0%; Score 10; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIVEDAE 10
DB 231 TACTIVEDAE 240

RESULT 5
Q9Y5X6 PRELIMINARY; PRT; 541 AA.
ID Q9Y5X6:
AC Q9Y5X6:
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Blood plasma glutamate carboxypeptidase precursor (EC 3.4.17.21).
GN PCGP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99223495; PubMed=10206990;

```


RA Gingras R., Richard C., El-Alfy M., Morales C.R., Potier M.,
 RA Feheretsky A.V.;
 RT "Purification, cDNA cloning, and expression of a new human blood
 RT plasma glutamate carboxypeptidase homologous to N-acetyl-aspartyl-
 RT alpha-glutamate carboxypeptidase/prostate-specific membrane antigen";
 RL J. Biol. Chem. 274:11742-11750 (1999).
 DR EMBL; AF119386; AAD31418.1; -.
 DR MEROPS; M28.014; -.
 DR InterPro; IPR003137; PA.
 DR PROSITE; PS50840; PA: 1.
 KW Carboxypeptidase; Hydrolase; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 45 541 BLOOD PLASMA GLUTAMATE CARBOXYPEPTIDASE.
 SQ SEQUENCE 541 AA; 59931 MW; 9B138FF897081F6B CRC64;
 QY Query Match 100.0%; Score 10; DB 4; Length 541;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 231 TACTIVEDAE 240
 RESULT 6
 Q8BGT9 PRELIMINARY; PRT; 576 AA.
 AC Q8BGT9;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Hypothetical glycosyl transferase.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cecum, and Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK033638; BAC28401.1; -.
 DR EMBL; AK042133; BAC31179.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 576 AA; 66540 MW; 706310A99FCDACD5 CRC64;
 QY Query Match 70.0%; Score 7; DB 11; Length 576;
 Best Local Similarity 100.0%; Pred. No. 7.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 500 ACTIVED 506
 RESULT 7
 Q9SM31 PRELIMINARY; PRT; 157 AA.
 AC Q9SM31;
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
 DE Hypothetical 18.1 kDa protein.
 GN F13M23.90 OR ATAG24950.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxId=3702;

RN (1)
 RP SEQUENCE FROM N.A.
 RA Bevan M., Wedler H., Wedler E., Wambutt R., Hoheisel J., Mewes H.W.,
 RA Mayer K.F.X., Scheller C.;
 RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RA Wedler H., Wedler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL035523; CAB36737.1; -.
 DR EMBL; AL161562; CAB79404.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 157 AA; 18115 MW; 72AFC01DA157BC3D CRC64;
 QY Query Match 60.0%; Score 6; DB 10; Length 157;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 39 ITVEDA 44
 RESULT 8
 Q9L990 PRELIMINARY; PRT; 167 AA.
 AC Q9L990;
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Vco28.
 GN VCO28.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrrio.
 OX NCBI_TaxId=666;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=569b;
 RX MEDLINE=20476433; PubMed=11021935;
 RA Clark C.A., Purins L., Kaewrakon P., Focareta T., Manning P.A.;
 RT "The Vibrio cholerae O1 chromosomal integrin";
 RL Microbiology 146:2605-2612 (2000).
 DR EMBL; AF179596; AAF71183.1; -.
 DR InterPro; IPR000182; GCSNacetyltransf.
 DR Pfam; PF00583; Acetyltransf. 1.
 SQ SEQUENCE 167 AA; 19002 MW; 559971F39B1D02F9 CRC64;
 QY Query Match 60.0%; Score 6; DB 2; Length 167;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 7 TVEDAE 12
 RESULT 9
 Q9KMC6 PRELIMINARY; PRT; 171 AA.
 AC Q9KMC6;
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Acetyltransferase, putative.

```

GN VCA0436.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Hatt D.H., Hickey E.K., Peterson J.D., Umayam U.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettein H., Richardson D.,
RA Ervolaeva M.D., Vanachavan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR TIGR; VCA0436; -; AAF96342.1; -.
DR InterPro; IPR000182; GCN5acetyltransf.
DR Pfam; PF00583; Acetyltransf; 1.
KW Transferrase; Complete proteome.
SQ SEQUENCE 171 AA; 19504 MW; CB18644753E0F834 CRC64;

Query Match 60.0%; Score 6; DB 16; Length 171;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ITVEDA 10
DB 11 ITVEDA 16

RESULT 10
Q8KV6 PRELIMINARY; PRT; 185 AA.
ID Q8KV6;
AC Q8KV6;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE 60 kDa chaparonin (Fragment).
GN CPN60.
OS uncultured pig faeces bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=190966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22034982; PubMed=12039767;
RA Hill J.E., Seipp R.P., Betts M., Hawkins L., Van Kessel A.G.,
RA Crosby W.L., Hemmingsen S.M.;
RT "Extensive profiling of a complex microbial community by high-
RT throughput sequencing.";
RL Appl. Environ. Microbiol. 68:3055-3066(2002).
DR EMBL; AF436976; AAM49237.1; -.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR Pfam; PF00118; cpn60_TCP1; 1.
KW ATP-binding; Chaparone.
FT NON_TER 1
FT NON_TER 185
SQ SEQUENCE 185 AA; 20485 MW; E9A9FC0397C5D322 CRC64;

Query Match 60.0%; Score 6; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
DB 85 ITVEDA 90

RESULT 11

```

```

Q8KVCO PRELIMINARY; PRT; 186 AA.
ID Q8KVCO;
AC Q8KVCO;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE 60 kDa chaparonin (Fragment).
GN CPN60.
OS uncultured pig faeces bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=190966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22034982; PubMed=12039767;
RA Hill J.E., Seipp R.P., Betts M., Hawkins L., Van Kessel A.G.,
RA Crosby W.L., Hemmingsen S.M.;
RT "Extensive profiling of a complex microbial community by high-
RT throughput sequencing.";
RL Appl. Environ. Microbiol. 68:3055-3066(2002).
DR EMBL; AF437097; AAM49358.1; -.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR Pfam; PF00118; cpn60_TCP1; 1.
KW ATP-binding; Chaparone.
FT NON_TER 1
FT NON_TER 186
SQ SEQUENCE 186 AA; 20188 MW; D798608CB07C9E7C CRC64;

Query Match 60.0%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
DB 85 ITVEDA 90

RESULT 12
Q8KV2 PRELIMINARY; PRT; 186 AA.
ID Q8KV2;
AC Q8KV2;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE 60 kDa chaparonin (Fragment).
GN CPN60.
OS uncultured pig faeces bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=190966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22034982; PubMed=12039767;
RA Hill J.E., Seipp R.P., Betts M., Hawkins L., Van Kessel A.G.,
RA Crosby W.L., Hemmingsen S.M.;
RT "Extensive profiling of a complex microbial community by high-
RT throughput sequencing.";
RL Appl. Environ. Microbiol. 68:3055-3066(2002).
DR EMBL; AF437042; AAM49303.1; -.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR Pfam; PF00118; cpn60_TCP1; 1.
KW ATP-binding; Chaparone.
FT NON_TER 1
FT NON_TER 186
SQ SEQUENCE 186 AA; 20151 MW; 7B5B5B5329114D CRC64;

Query Match 60.0%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
DB 85 ITVEDA 90

```

RESULT 13

Q8KH08

PRELIMINARY; PRT; 186 AA.

ID 08KH08

DT 01-OCT-2002 (TREMblrel. 22, Created)

DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)

DE 60 kDa chaperonin (Fragment).

GN CPN60.

OS unclutured pig faeces bacterium.

OC Bacteria; environmental samples.

OX NCBI_TaxID=190966;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22034982; PubMed=12039767;

RA Hill J.E., Seipp R.P., Betts M., Hawkins L., Van Kessel A.G.,

RA Crosby M.L., Hemmingsen S.M.,

RT "Extensive profiling of a complex microbial community by high-

RT throughput sequencing."

RL Appl. Environ. Microbiol. 68:3055-3066(2002).

DR EMBL; AF436897; AAA49158.1; -

DR EMBL; AF437007; AAA49258.1; -

DR EMBL; AF437129; AAA49390.1; -

DR InterPro; IPR002423; CPN60/TCP-1.

DR Pfam; PF00118; cpn60_TCP1; 1.

KM ATP-binding; Chaperone.

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 186 AA; 20201 MW; 7BB9B65B26290CEF CRC64;

Query Match 60.0%; Score 6; DB 2; Length 186;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 85 ITVEDA 90

QY 4 ITVEDA 9

ID 09PST2

AC 09PST2

DT 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)

DE RNA binding/signal transduction protein Qxi-4.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Hubbard White Mountain;

RX MEDLINE=98208728; PubMed=9547512;

RA Mezquita J., Pau M., Mezquita C.,

RT "Four isoforms of the signal-transduction and RNA-binding protein Qxi

RT expressed during chicken spermatogenesis."

RL Mol. Reprod. Dev. 50:70-78(1998).

CC -1- SIMILARITY: CONTAINS 1 KH DOMAIN.

DR EMBL; U86767; AAD06624.1; -

DR InterPro; IPR004087; KH_dom.

DR Pfam; PF00013; KH; 1.

DR SMART; SMO0322; KH; 1.

SQ SEQUENCE 186 AA; 21303 MW; A36DEB0C164CABC1 CRC64;

Query Match 60.0%; Score 6; DB 13; Length 186;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 ITVEDA 9

QY 4 ITVEDA 9

Db 155 ITVEDA 160

RESULT 15

Q9AB28

PRELIMINARY; PRT; 187 AA.

ID Q9AB28

DT 01-JUN-2001 (TREMblrel. 17, Created)

DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)

DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Hypothetical protein CC0406.

GN CC0406.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;

OC Caulobacteraceae; Caulobacter.

OX NCBI_TaxID=155892;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=ATCC 19089 / CB15;

RC MEDLINE=21173698; PubMed=11259647;

RA Nieman W.C., Feldlyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,

RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,

RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,

RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,

RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,

RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,

RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,

RT "Complete genome sequence of Caulobacter crescentus."

RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

DR EMBL; AE005713; AAK2393.1; -

DR TIGR; CC0406; -

KM Hypothetical protein; Complete proteome.

SQ SEQUENCE 187 AA; 20381 MW; F9617F6BEE2FCB85 CRC64;

Query Match 60.0%; Score 6; DB 16; Length 187;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 87 ITVEDA 92

QY 4 ITVEDA 9

ID 09PST2

AC 09PST2

DT 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)

DE RNA binding/signal transduction protein Qxi-4.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Hubbard White Mountain;

RX MEDLINE=98208728; PubMed=9547512;

RA Mezquita J., Pau M., Mezquita C.,

RT "Four isoforms of the signal-transduction and RNA-binding protein Qxi

RT expressed during chicken spermatogenesis."

RL Mol. Reprod. Dev. 50:70-78(1998).

CC -1- SIMILARITY: CONTAINS 1 KH DOMAIN.

DR EMBL; U86767; AAD06624.1; -

DR InterPro; IPR004087; KH_dom.

DR Pfam; PF00013; KH; 1.

DR SMART; SMO0322; KH; 1.

SQ SEQUENCE 186 AA; 21303 MW; A36DEB0C164CABC1 CRC64;

Query Match 60.0%; Score 6; DB 13; Length 186;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 ITVEDA 9

QY 4 ITVEDA 9

Search completed: December 22, 2003, 11:57:17

Job time : 37 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:55:16 ; Search time 22 Seconds
(without alignments)
19.232 Million cell updates/sec

Title: US-09-745-763-36_COPY_231_240

Perfect score: 10
Sequence: 1 TACTIVEDAE 10

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: Issued Patents AA.*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/6C.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	100.0	473	4 US-09-482-273-120	Sequence 120, App
2	6	60.0	351	4 US-09-252-991A-31251	Sequence 31251, A
3	6	60.0	371	2 US-08-837-593-8	Sequence 8, Appli
4	6	60.0	387	2 US-08-837-593-4	Sequence 4, Appli
5	6	60.0	393	2 US-08-837-593-3	Sequence 3, Appli
6	6	60.0	532	2 US-08-579-777A-2	Sequence 2, Appli
7	5	50.0	35	4 US-09-152-060-112	Sequence 112, App
8	5	50.0	98	4 US-09-511-881A-14	Sequence 14, Appl
9	5	50.0	103	6 5171684-5	Patent No. 5171684
10	5	50.0	115	4 US-09-252-991A-27427	Sequence 27427, A
11	5	50.0	129	3 US-09-011-074-2	Sequence 2, Appli
12	5	50.0	139	4 US-09-015-189-4	Sequence 4, Appli
13	5	50.0	145	3 US-09-011-074-6	Sequence 6, Appli
14	5	50.0	150	3 US-09-239-909-2	Sequence 2, Appli
15	5	50.0	158	4 US-09-663-600A-104	Sequence 104, App
16	5	50.0	180	4 US-09-252-991A-28297	Sequence 28297, A
17	5	50.0	185	4 US-09-107-532A-6934	Sequence 6934, Ap
18	5	50.0	186	4 US-09-328-352-6232	Sequence 6232, Ap
19	5	50.0	204	3 US-09-328-352-7175	Sequence 7175, Ap
20	5	50.0	224	3 US-09-091-889-10	Sequence 10, Appl
21	5	50.0	247	4 US-09-679-279-19	Sequence 19, Appl
22	5	50.0	264	4 US-09-328-352-6490	Sequence 6490, Ap
23	5	50.0	282	4 US-09-328-352-8229	Sequence 8229, Ap
24	5	50.0	293	4 US-09-252-991A-22215	Sequence 22215, A
25	5	50.0	294	4 US-09-674-741-6	Sequence 6, Appli
26	5	50.0	295	1 US-09-570-842-2	Sequence 2, Appli
27	5	50.0	295	1 US-08-700-576-2	Sequence 2, Appli

28	5	50.0	297	4 US-09-252-991A-28307	Sequence 28307, A
29	5	50.0	318	4 US-09-134-001C-4156	Sequence 4156, Ap
30	5	50.0	319	4 US-09-724-510-2	Sequence 2, Appli
31	5	50.0	319	4 US-09-723-216-2	Sequence 2, Appli
32	5	50.0	319	4 US-09-675-227-2	Sequence 2, Appli
33	5	50.0	328	4 US-08-828-242-1	Sequence 1, Appli
34	5	50.0	328	3 US-08-910-9278-1	Sequence 1, Appli
35	5	50.0	328	3 US-09-206-499-1	Sequence 1, Appli
36	5	50.0	328	3 US-09-270-370-1	Sequence 1, Appli
37	5	50.0	332	4 US-09-134-001C-4323	Sequence 4323, Ap
38	5	50.0	341	4 US-09-107-532A-7087	Sequence 7087, Ap
39	5	50.0	343	4 US-09-967-908A-8	Sequence 8, Appli
40	5	50.0	351	4 US-08-311-731A-28	Sequence 28, Appl
41	5	50.0	362	4 US-09-328-352-7264	Sequence 7264, Ap
42	5	50.0	363	4 US-09-967-908A-10	Sequence 10, Appl
43	5	50.0	372	4 US-09-967-908A-4	Sequence 4, Appli
44	5	50.0	375	2 US-08-506-864A-1	Sequence 1, Appli
45	5	50.0	375	2 US-08-851-968-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-482-273-120
; Sequence 120, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; EARLIER FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 120
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (473)
; OTHER INFORMATION: Xaa equals stop translation
US-09-482-273-120
Query Match 100.0%; Score 10; DB 4; Length 473;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 10; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
OY 1 TACTIVEDAE 10
Db 231 TACTIVEDAE 240
RESULT 2
US-09-252-991A-31251
; Sequence 31251, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31251
LENGTH: 351
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31251

Query Match 60.0%; Score 6; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9
DB 89 ITVEDA 94

RESULT 3
US-08-837-593-8
Sequence 8, Application US/08837593
Patent No. 5977442
GENERAL INFORMATION:
APPLICANT: Kleesig, Daniel F.
APPLICANT: Zhang Zhunqun
TITLE OF INVENTION: "No. 5977442el Salicylic Acid Induced
TITLE OF INVENTION: Map Kinase and Its Use for Enhanced Disease Resistance in Plant
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman,
STREET: 1601 Market Street, Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: United States of America
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/837,593
FILING DATE: April 21, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/029,805
FILING DATE: October 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
REFERENCE/DOCKET NUMBER: Rutgers 97-0016
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 371 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-837-593-8
Query Match 60.0%; Score 6; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 22;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 ITVEDA 9
DB 313 ITVEDA 318

RESULT 4
US-08-837-593-4
Sequence 4, Application US/08837593
Patent No. 5977442
GENERAL INFORMATION:
APPLICANT: Kleesig, Daniel F.
APPLICANT: Zhang Zhunqun
TITLE OF INVENTION: "No. 5977442el Salicylic Acid Induced
TITLE OF INVENTION: Map Kinase and Its Use for Enhanced Disease Resistance in Plant
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman,
STREET: 1601 Market Street, Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: United States of America
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/837,593
FILING DATE: April 21, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/029,805
FILING DATE: October 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
REFERENCE/DOCKET NUMBER: Rutgers 97-0016
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-837-593-4
Query Match 60.0%; Score 6; DB 2; Length 387;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 ITVEDA 9
DB 329 ITVEDA 334
RESULT 5
US-08-837-593-3
Sequence 3, Application US/08837593
Patent No. 5977442
GENERAL INFORMATION:
APPLICANT: Kleesig, Daniel F.

APPLICANT: Zhang Zhugun
TITLE OF INVENTION: "No. 5977442el Salicylic Acid Induced
TITLE OF INVENTION: Map Kinase and its Use for Enhanced Disease Resistance in Plant
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS: 14
ADDRESSEE: Dann, Dorfman, Herrell and Skillman,
ADDRESSEE: P.C.
STREET: 1601 Market Street, Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: United States of America
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/837,593
FILING DATE: April 21, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/029,805
FILING DATE: October 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
REFERENCE/DOCKET NUMBER: Rutgers 97-0016
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
US-08-837-593-3

Query Match 60.0%; Score 6; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9
Db 334 ITVEDA 339

RESULT 6
US-08-579-777A-2
Sequence 2, Application US/08579777A
Patent No. 5912153
GENERAL INFORMATION:
APPLICANT: Enderlin, Carol S.
APPLICANT: Seliterminkoff, Claude P.
TITLE OF INVENTION: (1,3)-B-glucan synthase genes and
TITLE OF INVENTION: inducible inhibition of fungal growth using the antisense
TITLE OF INVENTION: constructs derived therefrom.
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Davis, Graham & Stubbs L.L.P.
STREET: 370 Seventeenth Street
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80201-0185
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/579,777A
FILING DATE: 28-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/155,004
FILING DATE: 18-NOV-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 892-7370
TELEFAX: (303) 893-1379
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neurospora crassa
STRAIN: wild-type 74-DR 23-1VA
DEVELOPMENTAL STAGE: Mycelial

US-08-579-777A-2

Query Match 60.0%; Score 6; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TVEDAE 10
Db 520 TVEDAE 525

RESULT 7
US-09-152-060-112
Sequence 112, Application US/09152060
Patent No. 6448230
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 28 Human Secreted Proteins
FILE REFERENCE: P2003Pl.US
CURRENT APPLICATION NUMBER: US/09/152,060
CURRENT FILING DATE: 1998-09-11
EARLIER APPLICATION NUMBER: PCT/US98/04858
EARLIER FILING DATE: 1998-03-12
EARLIER APPLICATION NUMBER: 60/040,762
EARLIER FILING DATE: 1997-03-14
EARLIER APPLICATION NUMBER: 60/040,710
EARLIER FILING DATE: 1997-03-14
EARLIER APPLICATION NUMBER: 60/050,934
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/048,100
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/048,357
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/048,189
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/057,765
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/068,368
EARLIER FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 118
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 112
LENGTH: 35
TYPE: PRT

ORGANISM: Homo sapiens
US-09-152-060-112

Query Match 50.0%; Score 5; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVED 8
|||
Db 6 ITVED 10

RESULT 8
US-09-511-881A-14
Sequence 14, Application US/09511881A
Patent No. 6495674

GENERAL INFORMATION:
APPLICANT: SALK INSTITUTE
APPLICANT: LEMKE, GREGORY
APPLICANT: NGUYEN, ANDREW
APPLICANT: KRAPP, RALF

TITLE OF INVENTION: ELECTINS AND THEIR USE
FILE REFERENCE: SALIKINS.001A
CURRENT APPLICATION NUMBER: US/09/511,881A
CURRENT FILING DATE: 2000-02-25

NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 96
TYPE: PRT

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN

LOCATION: (1) --(98)
OTHER INFORMATION: Pleckstrin homology domain
US-09-511-881A-14

Query Match 50.0%; Score 5; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVED 8
|||
Db 56 ITVED 60

RESULT 9
5171684-5

PATENT NO. 5171684
APPLICANT: YEN, KWANG-MU; BLATT, LAWRENCE M.; KARL, MICHAEL R.
TITLE OF INVENTION: BIOCONVERSIONS CATALYZED BY THE TOULUENE
MONOOXYGENASE OF PSEUDOMONAS MENDOCINA KR-1

NUMBER OF SEQUENCES: 41
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/590,374
FILING DATE: 28-SEP-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 177,631
FILING DATE: 05-APR-1988
SEQ ID NO: 5
LENGTH: 103

Query Match 50.0%; Score 5; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVED 8
|||
Db 40 ITVED 44

RESULT 10

US-09-252-991A-27427
Sequence 27427, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27427
LENGTH: 115
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27427

Query Match 50.0%; Score 5; DB 4; Length 115;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 VEDAE 10
|||
Db 109 VEDAE 113

RESULT 11
US-09-011-074-2
Sequence 2, Application US/09011074
Patent No. 6133005

GENERAL INFORMATION:
APPLICANT: Pouska, Annemarie
APPLICANT: Coy, Johannes

TITLE OF INVENTION: TRANSKETOLASE-RELATED PROTEIN
FILE REFERENCE: 8484-0039-999
CURRENT APPLICATION NUMBER: US/09/011,074
CURRENT FILING DATE: 1998-05-21

EARLIER APPLICATION NUMBER: PCT/DE96/01401
EARLIER FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 129
TYPE: PRT

ORGANISM: Homo sapien
US-09-011-074-2

Query Match 50.0%; Score 5; DB 3; Length 129;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVED 8
|||
Db 64 ITVED 68

RESULT 12
US-09-015-189-4

Sequence 4, Application US/09015189A
Patent No. 6410263
GENERAL INFORMATION:
APPLICANT: Wallis, Nicola G.

TITLE OF INVENTION: No. 6410263el Histidine Kinase
FILE REFERENCE: GMS0008
CURRENT APPLICATION NUMBER: US/09/015,189A
CURRENT FILING DATE: 1998-01-29

EARLIER APPLICATION NUMBER: 60/037,856
EARLIER FILING DATE: 1997-02-07
NUMBER OF SEQ ID NOS: 6


```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-015-189-4
```

```
Query Match          50.0%; Score 5; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 VEDAE 10
      |||||
Db      69 VEDAE 73
```

```
RESULT 13
US-09-011-074-6
; Sequence 6, Application US/09011074
; Patent No. 6133005
; GENERAL INFORMATION:
; APPLICANT: Poustka, Annemarie
; APPLICANT: Coy, Johannes
; TITLE OF INVENTION: TRANSKETOLASE-RELATED PROTEIN
; FILE REFERENCE: 8484-0039-999
; CURRENT APPLICATION NUMBER: US/09/011,074
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: PCT/DE96/01401
; EARLIER FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-011-074-6
```

```
Query Match          50.0%; Score 5; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 ITVED 8
      |||||
Db      80 ITVED 84
```

```
RESULT 14
US-09-239-909-2
; Sequence 2, Application US/09239909
; Patent No. 6284952
; GENERAL INFORMATION:
; APPLICANT: Kumo Petrochemical Co. Ltd.
; TITLE OF INVENTION: Transgenic Plants with Divergent SCAM4 or SCAM5 Gene to Achieve M
; FILE REFERENCE: P99-2-6
; CURRENT APPLICATION NUMBER: US/09/239,909
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: EP 99300136.1
; EARLIER FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 2
; LENGTH: 150
; TYPE: PRT
; ORGANISM: G. max calmodulin4 (SCAM4)
US-09-239-909-2
```

```
Query Match          50.0%; Score 5; DB 3; Length 150;
Best Local Similarity 100.0%; Pred. No. 1,2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 CITVE 7
      |||||
Db      27 CITVE 31
```

```
RESULT 15
US-09-663-600A-104
; Sequence 104, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclet, Aymeric
; APPLICANT: Bouguetel, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent .pm
; SEQ ID NO 104
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -37...-1
US-09-663-600A-104
```

```
Query Match          50.0%; Score 5; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 1,2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 ACTIV 6
      |||||
Db      37 ACTIV 41
```

Search completed: December 22, 2003, 11:58:24
Job time : 22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:56:11 ; Search time 30 Seconds
(without alignments)
62.253 Million cell updates/sec

Title: US-09-745-763-36_COPY_231_240
Perfect score: 10
Sequence: 1 TACTIVEDAR 10

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 696363 seqs, 186758610 residues

Word size : 0

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.0	472	9	US-09-745-763-36	Sequence 36, Appl
2	100.0	473	11	US-09-984-271-120	Sequence 120, App
3	60.0	432	9	US-09-815-242-11700	Sequence 11700, A
4	60.0	433	12	US-09-946-290-14	Sequence 14, Appl
5	60.0	610	9	US-09-815-242-5559	Sequence 5559, Ap
6	60.0	618	9	US-09-815-242-12567	Sequence 12567, A
7	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
8	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
9	60.0	1114	14	US-09-771-161A-163	Sequence 163, Appl
10	60.0	1115	10	US-09-771-161A-253	Sequence 253, Appl
11	60.0	1115	10	US-10-350-405-201	Sequence 201, Appl
12	60.0	1115	10	US-10-350-405-201	Sequence 201, Appl
13	50.0	18	15	US-10-057-789-239	Sequence 239, App
14	50.0	18	15	US-10-212-628-239	Sequence 239, App
15	50.0	35	9	US-09-853-161-112	Sequence 112, App

16	5	50.0	35	9	US-09-852-659A-112	Sequence 112, App
17	5	50.0	35	10	US-09-852-797-112	Sequence 112, App
18	5	50.0	50	12	US-10-371-634-26	Sequence 26, Appl
19	5	50.0	55	9	US-09-864-761-34650	Sequence 34650, A
20	5	50.0	60	12	US-10-029-386-29649	Sequence 29649, A
21	5	50.0	94	15	US-10-204-887-114	Sequence 114, App
22	5	50.0	104	12	US-10-177-478-38	Sequence 38, Appl
23	5	50.0	104	15	US-10-156-761-8337	Sequence 8337, Ap
24	5	50.0	128	12	US-10-029-386-32159	Sequence 32159, A
25	5	50.0	129	10	US-09-862-658-5	Sequence 5, Appl
26	5	50.0	129	15	US-10-051-767-11	Sequence 11, Appl
27	5	50.0	129	15	US-10-175-696-26	Sequence 26, Appl
28	5	50.0	132	10	US-09-738-626-4098	Sequence 4098, Ap
29	5	50.0	132	15	US-10-156-761-12474	Sequence 12474, A
30	5	50.0	150	9	US-09-925-301-1372	Sequence 1372, Ap
31	5	50.0	158	12	US-10-319-763-104	Sequence 104, App
32	5	50.0	176	9	US-09-780-177-2	Sequence 2, Appl
33	5	50.0	184	10	US-09-984-245-218	Sequence 218, App
34	5	50.0	184	11	US-09-966-262-218	Sequence 218, App
35	5	50.0	184	11	US-09-983-966-218	Sequence 218, App
36	5	50.0	184	15	US-10-143-090-218	Sequence 218, App
37	5	50.0	190	10	US-09-738-626-4637	Sequence 4637, Ap
38	5	50.0	191	12	US-09-981-151A-83	Sequence 83, Appl
39	5	50.0	198	10	US-09-925-300-1199	Sequence 1199, Ap
40	5	50.0	200	11	US-09-764-891-4187	Sequence 4187, Ap
41	5	50.0	210	16	US-10-080-170-58	Sequence 58, Appl
42	5	50.0	217	10	US-09-738-626-5801	Sequence 5801, A
43	5	50.0	228	12	US-10-029-386-32350	Sequence 32350, A
44	5	50.0	240	15	US-10-106-698-6361	Sequence 6361, Ap
45	5	50.0	244	9	US-09-867-550-1200	Sequence 1200, Ap

ALIGNMENTS

RESULT 1
US-09-745-763-36
Sequence 36, Application US/09745763
Patent No. US20020065394A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaValle, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Merberg, David
Treacy, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-745-763-36

Query Match 100.0%; Score 10; DB 9; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIVEDAE 10
|||||
DB 231 TACTIVEDAE 240

RESULT 2

US-09-984-271-120
Sequence 120, Application US/09984271
Publication No. US20030040088A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/984,271
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: 09/482,273
PRIOR FILING DATE: 2000-01-13
PRIOR APPLICATION NUMBER: PCT/US99/15849
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: 60/092,921
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/092,922
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/092,956
PRIOR FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 120
LENGTH: 473
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (473)
OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-120

Query Match 100.0%; Score 10; DB 11; Length 473;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIVEDAE 10
|||||
DB 231 TACTIVEDAE 240

RESULT 3

US-09-815-242-11700
Sequence 11700, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11700
LENGTH: 432
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-815-242-11700

Query Match 60.0%; Score 6; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
|||||
DB 89 ITVEDA 94

RESULT 4

US-09-946-290-14
Sequence 14, Application US/09946290
Publication No. US2003021475A1
GENERAL INFORMATION:
APPLICANT: Roberts, Christopher J.
TITLE OF INVENTION: METHODS FOR IDENTIFYING PATHWAY-SPECIFIC REPORTERS AND
FILE REFERENCE: 9301-040
CURRENT APPLICATION NUMBER: US/09/946,290
CURRENT FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: US/09/282,243
PRIOR FILING DATE: 1999-03-31
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 433
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-946-290-14

Query Match 60.0%; Score 6; DB 12; Length 433;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
|||||
DB 307 ITVEDA 312

RESULT 5

US-09-815-242-5559
Sequence 5559, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5559
LENGTH: 610
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5559

Query Match 60.0%; Score 6; DB 9; Length 610;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
Db 595 TVEDAE 600

RESULT 6
US-09-815-242-12567
Sequence 12567, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12567
LENGTH: 618
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12567

Query Match 60.0%; Score 6; DB 9; Length 618;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
Db 603 TVEDAE 608

RESULT 7
US-09-815-242-12970
Sequence 12970, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12970
LENGTH: 618
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12970

Query Match 60.0%; Score 6; DB 9; Length 618;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10
Db 603 TVEDAE 608

RESULT 8
US-09-771-161A-163
Sequence 163, Application US/09771161A
Patent No. US2002011081A1
GENERAL INFORMATION:
APPLICANT: LEVINE, et al.

```

; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 163
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-771-161A-163

Query Match          60.0%; Score 6; DB 10; Length 842;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TWEEDAE 10
   |||||
   70 TWEEDAE 75

Db

RESULT 9
US-10-005-983-2
; Sequence 2, Application US/10005983
; Publication No. US20020116730A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Keith D
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING PERK PROTEIN
; FILE REFERENCE: R-517
; CURRENT APPLICATION NUMBER: US/10/005,983
; CURRENT FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,676
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 60/311,018
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/324,765
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/326,148
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1114
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-005-983-2

Query Match          60.0%; Score 6; DB 14; Length 1114;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TWEEDAE 10
   |||||
   67 TWEEDAE 72

Db

RESULT 10
US-09-771-161A-253
; Sequence 253, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
```

```

; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 253
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-771-161A-253

Query Match          60.0%; Score 6; DB 10; Length 1115;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TWEEDAE 10
   |||||
   70 TWEEDAE 75

Db

RESULT 11
US-10-296-540-2
; Sequence 2, Application US/10296540
; Publication No. US20030215827A1
; GENERAL INFORMATION:
; APPLICANT: JULIER, C, cile
; APPLICANT: DELEPINE, Marc
; APPLICANT: NICOLINO, Marc
; TITLE OF INVENTION: MUTATED EUKARIOTIC TRANSLATION INITIATION FACTOR 2 ALPHA KINASE 3,
; TITLE OF INVENTION: MULTIPLE EPIPHYSEAL DYSPLASIA (WOLCOTT-RALLISON SYNDROME)
; FILE REFERENCE: 344 061 - US
; CURRENT APPLICATION NUMBER: US/10/296,540
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: EP 00/401 436
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: EP 00/402 707
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: PCT/IB 01/01 153
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-296-540-2

Query Match          60.0%; Score 6; DB 12; Length 1115;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TWEEDAE 10
   |||||
   70 TWEEDAE 75

Db

RESULT 12
US-10-350-405-201
; Sequence 201, Application US/10350405
; Publication No. US20030215894A1
; GENERAL INFORMATION:
; APPLICANT: Niman, Henry L.
; TITLE OF INVENTION: Polypeptide-induced Monoclonal Receptors
; TITLE OF INVENTION: to Protein Ligands
; FILE REFERENCE: TSRI 35.5 CON 7/LIG
; CURRENT APPLICATION NUMBER: US/10/350,405
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: 09/427,576
; PRIOR FILING DATE: 1999-10-26
```

;; PRIOR APPLICATION NUMBER: 08/461,563
;; PRIOR FILING DATE: 1995-06-02
;; PRIOR APPLICATION NUMBER: 08/294,879
;; PRIOR FILING DATE: 1994-08-23
;; PRIOR APPLICATION NUMBER: 08/054,864
;; PRIOR FILING DATE: 1993-04-28
;; PRIOR APPLICATION NUMBER: 07/900,502
;; PRIOR FILING DATE: 1992-06-16
;; PRIOR APPLICATION NUMBER: 07/780,415
;; PRIOR FILING DATE: 1991-10-22
;; NUMBER OF SEQ ID NOS: 227
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 201
;; LENGTH: 12
;; TYPE: PRT
;; ORGANISM: retrovirus
US-10-350-405-201

Query Match 50.0%; Score 5; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CITE 7
|||
Db 1 CITE 5

RESULT 13
US-10-057-789-239

;; Sequence 239, Application US/10057789
;; Publication No. US20030082522A1
;; GENERAL INFORMATION:
;; APPLICANT: Paul Haynes
;; APPLICANT: Jing Wei
;; APPLICANT: John Yates
;; APPLICANT: Nancy Andon
;; TITLE OF INVENTION: DIFFERENTIAL LABELING FOR QUANTITATIVE
;; TITLE OF INVENTION: ANALYSIS OF COMPLEX PROTEIN MIXTURES
;; FILE REFERENCE: NADII.022A
;; CURRENT APPLICATION NUMBER: US/10/057,789
;; CURRENT FILING DATE: 2002-06-28
;; PRIOR APPLICATION NUMBER: US 60/264,576
;; PRIOR FILING DATE: 2001-01-26
;; PRIOR APPLICATION NUMBER: US 60/305,232
;; PRIOR FILING DATE: 2001-07-13
;; NUMBER OF SEQ ID NOS: 311
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 239
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Saccharomyces cerevisiae
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: 4
;; OTHER INFORMATION: Xaa = Modified Cysteine
US-10-057-789-239

Query Match 50.0%; Score 5; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 VEDAE 10
|||
Db 11 VEDAE 15

RESULT 14
US-10-212-628-239

;; Sequence 239, Application US/10212628
;; Publication No. US20030087329A1
;; GENERAL INFORMATION:
;; APPLICANT: Paul Haynes
;; APPLICANT: Jing Wei

;; APPLICANT: John Yates
;; APPLICANT: Nancy Andon
;; TITLE OF INVENTION: DIFFERENTIAL LABELING FOR QUANTITATIVE
;; TITLE OF INVENTION: ANALYSIS OF COMPLEX PROTEIN MIXTURES
;; FILE REFERENCE: NADII.022CPI
;; CURRENT APPLICATION NUMBER: US/10/212,628
;; CURRENT FILING DATE: 2002-08-01
;; PRIOR APPLICATION NUMBER: US 60/264,576
;; PRIOR FILING DATE: 2001-01-26
;; PRIOR APPLICATION NUMBER: US 60/305,232
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: US 10/057,789
;; PRIOR FILING DATE: 2002-01-25
;; NUMBER OF SEQ ID NOS: 311
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 239
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Saccharomyces cerevisiae
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: 4
;; OTHER INFORMATION: Xaa = Modified Cysteine
US-10-212-628-239

Query Match 50.0%; Score 5; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 VEDAE 10
|||
Db 11 VEDAE 15

RESULT 15
US-09-853-161-112

;; Sequence 112, Application US/09853161
;; Patent No. US20020076756A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: 28 Human Secreted Proteins
;; FILE REFERENCE: P2003P3
;; CURRENT APPLICATION NUMBER: US/09/853,161
;; CURRENT FILING DATE: 2001-05-11
;; PRIOR APPLICATION NUMBER: 60/265,583
;; PRIOR FILING DATE: 2001-02-02
;; PRIOR APPLICATION NUMBER: 09/152,060
;; PRIOR FILING DATE: 1998-09-11
;; PRIOR APPLICATION NUMBER: PCT/US98/04858
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/040,762
;; PRIOR FILING DATE: 1997-03-14
;; PRIOR APPLICATION NUMBER: 60/040,710
;; PRIOR FILING DATE: 1997-03-14
;; PRIOR APPLICATION NUMBER: 60/050,934
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: 60/048,100
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: 60/048,357
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: 60/048,189
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: 60/057,765
;; PRIOR FILING DATE: 1997-09-05
;; PRIOR APPLICATION NUMBER: 60/048,970
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/068,368
;; PRIOR FILING DATE: 1997-12-19
;; NUMBER OF SEQ ID NOS: 118
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 112
;; LENGTH: 35
;; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-853-161-112

Query Match 50.0%; Score 5; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 ITVED 8
|||
Db 6 ITVED 10

Search completed: December 22, 2003, 11:59:08
Job time : 31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 22, 2003, 20:22:58 ; Search time 4346 Seconds
(without alignments)
4443.015 Million cell updates/sec

Title: US-09-745-763-36
Perfect score: 2417
Sequence: 1 MKFLIFAFGCVHLISLCSG.....AVNAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2886711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-Q=/cgn2_1/USPRO.spool/p/US09745763/runat_22122003_11317_25251/app_query.fasta.1.647
-DB=GenBml -QPMT=fastap -SUFFIX=g2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09745763 @CGN 1.1.3745 @runat_22122003_11317_25251 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenBml.*
1: gb_da.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_ats.*
28: em_un.*

29: em_vi.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pin.*
35: em_hcg_rod.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hcgo_hum.*
40: em_hcgo_mus.*
41: em_hcgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2417	100.0	1472	9	AF107834	AF107834 Homo sapi
2	2417	100.0	1851	6	BD106411	BD106411 Secreted
3	2417	100.0	1860	6	AX006440	AX006440 Sequence
4	2417	100.0	1928	9	BC020689	BC020689 Homo sapi
5	2410	99.7	1778	6	BD105816	BD105816 Secretory
6	2402	99.4	1860	6	BD127520	BD127520 Primer fo
7	2402	99.4	1860	9	AK075132	AK075132 Homo sapi
8	2384	98.6	1794	9	AF119386	AF119386 Homo sapi
9	2133	88.2	1596	10	BC037067	BC037067 Mus muscu
10	2133	88.2	1726	10	AF131077	AF131077 Rattus no
11	2117	87.6	1778	6	AX400829	AX400829 Sequence
12	2116	87.5	1778	10	AF097723	AF097723 Rattus no
13	2019	87.5	1778	10	AF097723	AF097723 Mus muscu
14	2029	83.9	1716	10	AF009513	BD059610 Secreted
15	870	36.0	609	6	BD059610	BD059610 Secreted
16	771	31.9	642	6	BD126476	BD126476 Primer fo
17	771	31.9	642	6	BD126476	BD126476 Primer fo
18	769.5	31.8	1729	3	AF077194	AF077194 Acanthoch
19	745	30.8	157927	9	AC084223	AC084223 Homo sapi
20	745	30.8	157927	9	AP006278	AP006278 Homo sapi
21	745	30.8	171301	9	AC010859	AC010859 Homo sapi
22	701	29.0	423	9	AF107833	AF107833 Homo sapi
23	662	27.4	10792	1	AE015790	AE015790 Shewanell
24	640	26.5	10487	1	AE005922	AE005922 Caulobact
25	623	25.8	417	10	AF107833	AF107833 Mus muscu
26	606	25.1	167883	2	AC121026	AC121026 Rattus no
27	606	25.1	218874	2	AC110103	AC110103 Rattus no
28	600	24.8	214287	1	AC129951	AC129951 Mus muscu
29	513	21.2	10927	1	AE012431	AE012431 Xanthomon
30	499.5	20.7	11712	1	AE011976	AE011976 Xanthomon
31	491	20.3	462	6	BD105811	BD105811 Secretory
32	491	20.3	462	6	E62849	E62849 Secretory p
33	479	19.8	314	6	BD058492	BD058492 Secreted
34	393	16.3	144703	9	AP003112	AP003112 Homo sapi
35	351	14.5	317	6	BD076934	BD076934 5' EST of
36	346	14.3	152176	9	AP003117	AP003117 Homo sapi
37	346	14.3	171035	9	AP003111	AP003111 Homo sapi
38	340	14.1	218874	2	AC110103	AC110103 Rattus no
39	334	13.8	214287	2	AC129951	AC129951 Mus muscu
40	332	13.7	203469	2	AC123654	AC123654 Mus muscu
41	302	12.5	238034	2	AC123681	AC123681 Mus muscu
42	298	12.3	166050	2	AC013817	AC013817 Homo sapi
43	298	12.3	177008	2	AC011175	AC011175 Homo sapi
44	297	12.3	240178	2	AC094173	AC094173 Rattus no
45	297	12.3	288136	2	AC120489	AC120489 Rattus no

RESULT 1

ALIGNMENTS

AF107834 1472 bp mRNA linear PRI 12-JUL-1999
 LOCUS Homo sapiens clone LCH1 aminopeptidase mRNA, complete cds.
 DEFINITION AF107834
 ACCESSION AF107834
 VERSION AF107834.1 GI:5442029
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1472)
 AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.
 TITLE Cloning of the human aminopeptidase gene
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 1472)
 AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.
 TITLE Direct Submission
 JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
 Sinica, RM 433, 128, Yen-Chun-Yuan Road SEC 2, Taipei 11529, Taiwan
 FEATURES
 source
 1..1472
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="LCH1"
 /tissue_type="liver"
 /dev_stage="fetus"
 54..1472
 /note="secretory protein"
 /codon_start=1
 /product="aminopeptidase"
 /protein_id="AAD43214.1"
 /db_xref="GI:5442030"
 /translation="MKELIPAFPGVNHLSGSKATKNGISKRTSEIKESIASCS
 DVAKAIINLAVYKAQNSRYERLALVDTPVRISGSNLEKAIQIMONQOGLK
 VHLSPVRLPHEWREBSAVMLEPRHKTAIILGLSSISCTPEGIITAEVLVTSDEIQ
 RRAEARKGIIVYNQPIYNGRTVOYRQGAVEAKVGLASLRSVASFISPHNG
 IOEYODGVKIPACITVEDAMMSRMAHGIKIYOLKMAKTYPDRTSPNTVAEIT
 GSKYDEYVLVSLGHLDSWDGQANDDGGAFLISWELSLIKDIGLRKRLVLT
 AEEGGVGAFOYQLHKVNIISYSLVMSDGTFLPTGLDTGSERKARLIEEWSLL
 QPLNTQVLSHGEGTDINFWIQAGVPGASLLDLIDYKFFPHSHGDTWMDPKOMV
 AAAYWAVSYVAWDEMLPRPS"
 CDS
 411 a 325 c 381 g 355 t
 BASE COUNT
 ORIGIN
 Alignment Scores:
 Pred. No.: 8,836-176 Length: 1472
 Score: 2417.00 Matches: 472
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-745-763-36 (1-472) x AF107834 (1-1472)
 Oy 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCyseSergly 20
 Db 54 ATGAATTCCTTATCTTGCATTTTCCGTGCTTCACTTTATCCCTGCTCTGGG 113
 Oy 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
 Db 114 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAATAAAGAAATA 173
 Oy 41 AlaSerCyGlyAspValAlaLysAlaIleIleLeuLeuAlaValTyrGlyLysAlaGln 60
 Db 174 GCCAGCTGTGGAGATGCTTAAGCAATCATCACTAGCTGTTATGTATTAAGCCAG 233
 Oy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSergly 80
 Db 234 AACGATCTATAGACGATGGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 293
 Oy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAlaAspGlyLeu 100

Db 294 TCCAAGAACCTAGAAAAAGCATCAATATTATGTAACCAAAAGCTGCAGCAAGATGGGCTG 353
 Oy 101 GluLysValHisLeuGluProValArgIleProHisIleProGluValArgGlyGluSerAla 120
 Db 354 GAGAAAGTCACTCGAGGAGGATGAGATACCCCATCTGGAGAGGAGGAGAAATCAAGCT 413
 Oy 121 ValMetLeuGluProValGlnIleLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
 Db 414 CTGATGCTGAGCCAAATAATTCATATAGATAGCCATCTCGGGTCTGGACAGCATTTGGG 473
 Oy 141 ThrProProGluGlyIleThrAlaGluValLeuValHisSerPheAspGluLeuGln 160
 Db 474 ACTCTCTCAGAGGATTAAGCAGCAAGAGTTCTGTGTGTGCTGCTCTTCTGATGAACTCAG 533
 Oy 161 ArgArgAlaSerGluAlaArgGlyValValValValTyrAsnGlnProTyrIleAsnTyr 180
 Db 534 AGAAGGCTCTAGAGCAAGAGGAGAAATGTTGTTATTAACCACTTACATCAATCACTAC 593
 Oy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200
 Db 594 TCAGAGACGGTGCATACCGAACCGAGGGGGCGGTGGAACTGCCAAGTTGGGCTTTG 653
 Oy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleGln 220
 Db 654 GCATCTCTCATTCGATCCGTGGCTCTCTCTCCATCTACAGTCTCTCACAGGCAATTCAG 713
 Oy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 Db 714 GAATACCAAGATGCGCGCCCAAGATTCAACAGCTGTATTCAGGAGAAAGCAGCA 773
 Oy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 Db 774 ATGATGTCAGGAATGCGCTTCTCATAGGATCAAAATTCATTCACGTAAAGATGGGGCA 833
 Oy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
 Db 834 AAGACCTACCAAGATCTGATTCCTTCAACACTGTACAGAGATCATCTGGAGCAAAATAT 893
 Oy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300
 Db 894 CCAGAACAGGTGTACTGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 953
 Oy 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrProGluAlaLeuSerLeuIleLysAspLeu 320
 Db 954 ATGATATATGCGCGGTGAGCTTATATCATGGAAACATCTCATTATTAAGATCTT 1013
 Oy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrThrAlaGluGluGlnGly 340
 Db 1014 GGGCTGCTGCCAAAGAGAGACTGCGGCTGTGCTGTGCACTGCAGAAAGCAAGGTGGA 1073
 Oy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
 Db 1074 GTTGTCCTTCAGATATACATGATACAGTACAGAGTAAATTTTCAACTACAGCTGAGTG 1133
 Oy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
 Db 1134 ATGAGCTGAGCGAGAGAACCTTTTACCACTGGGTGCAATTCATGCGCAAGTGAAG 1193
 Oy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 Db 1194 GCCAGGCGCATATGAGAGAGGTATAGCTGCTGCAGGCCCTTCATATACACTCAGGTC 1253
 Oy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420
 Db 1254 CTGAGCATGAGAGAGAGAGACATCACTTTTGATCCAAAGCTGAGTGCCTGAGGCC 1313
 Oy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisIleSerHisGlyAspThrMet 440
 Db 1314 AGTCTACTGATGATCTTATACAGATATTTCTTCTTCATCACTCCACGAGACACCATG 1373
 Oy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaAlaValThrAlaValSerTyr 460
 Db 1374 ACTGTATGATGCCAAGAGATGATGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTTAT 1433

Qy 461 ValValAlaAspMetGluMetLeuProArgSer 472
Db 1434 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1469

RESULT 2
LOCUS BD106411 1851 bp DNA linear PAT 18-SEP-2002
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD106411
VERSION BD106411.1 GI:23201229
KEYWORDS JP 2002503955-A/2.
SOURCE Chlamydia sp.
ORGANISM Chlamydia sp.
REFERENCE 1 (bases 1 to 1851)
AUTHORS Jacobo, K., McCoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D., Treacy, M., Spaulding, V., and Agostino, M.J.
TITLE Secreted proteins and polynucleotides encoding them
JOURNAL Patent: JP 2002503955-A 2 05-FEB-2002;
GENETICS INSTITUTE INC
COMMENT PN JP 2002503955-A/2
PD 05-FEB-2002
PR 20-MAR-1998 JP 1998545874
PR 20-MAR-1997 US 08/822167, 19-MAR-1998 US 09/044466 PI
KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
C12N15/12, C07K14/47, A61K8/17
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.

FEATURES
source 1..1851
/organism="Chlamydia sp."
/mol_type="genomic DNA"
/db_xref="taxon:35827"

BASE COUNT 531 a 413 c 438 g 469 t
ORIGIN

Alignment Scores:
Pred. No.: 1,18e-175 Length: 1851
Score: 2417.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 6

US-09-745-763-36 (1-472) x BD106411 (1-1851)

Qy 1 MetLysSheLeuLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20
Db 99 ATGAAATTCCTTACTTCCTGCGATTTTTCGCGTGTTCACCTTTATCCCTGCTCGGG 158

Qy 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
Db 159 AAACCTATATCCAGATGCGATCTCTAAGAGCATTTTGAAGAAATTAAGAAATA 218

Qy 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTYGlyLysAlaGln 60
Db 219 GCCAGCTGTGGAGATGTTGCTAAAGCAATCATCACTGCTTTATGTGTAAGCCAG 278

Qy 61 AsnArgSerTYRGLUArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db 279 AACGATTCCTAATGCGCATTCCTGCTGATATCTGTGAGCCCACTGAGTGGC 338

Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTYRGLNAsnLeuGlnAlaAspGlyLeu 100
Db 339 TCCAGAACTTACGAAAGAAAGCAATTCCTCAATTTATGACCAAACTGCGAGCAATGGGCTG 398

Qy 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluGluSerAla 120
Db 399 GAGAAAGTTCACCTGAGCCAGTGAATATCCCACTGAGAGGGGAGAGAAATCAGCT 458

Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 459 GTGATGTGGAGCCAGAAATTCATTAAGATAGCATCTGGGTCTTGGCAGCAGATTGGG 518

Qy 141 ThrProGluGluGlyIleThrAlaGluValLeuValHisSerPheAspGluLeuGln 160
Db 519 ACTCTTCAGAAAGCAATTCAGCAGAAAGTTCTGGTGGTGAACCTTTCATGTAATCGCAG 578

Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTYRAsnGlnProTYRlleAsnTYR 180
Db 579 AGAAGGGCTTCAGAAAGCAAGAGGAGATTGTTTATTAACCAACTTATCATCACTAC 638

Qy 181 SerArgThrValGlnTYRArgTrpGlnGlyAlaValGluAlaLysValGlyAlaLeu 200
Db 639 TCAAGACGATGCAATTCACCAAGCAGGAGGCGGTGAGACTCCCAAGTGGGGCTTTG 658

Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTYRSerProHisTrpGlyIleGln 220
Db 699 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCATGACGTCTTCACAGGATTCAG 758

Qy 221 GluTYRGLNAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 759 GAATACCAAGATGCGCGCCCAAAATTCACAGCCCTGTATTACGGTGAAGATGCAGAA 818

Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 819 ATGATGTCAGAAATGCGCTTCTCATGGATCAAAATTCCTTACCTTAAGATGGGGCA 878

Qy 261 LysThrTYRProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTYR 280
Db 879 AAGCTACCAAGATGCTGATTCCTTCAACACTGTAGCAAGATCATCTGGAGAAATAT 938

Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
Db 939 CCAAGACAGGTTGACTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 998

Qy 301 MetAspAspGlyGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
Db 999 ATGATGATGCGCGTGGAGCTTTTATTCATGGGACACTCTTCATTTAAGATCTT 1058

Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpHisAlaGluGlnGlyGly 340
Db 1059 GGGGTGGTCCAAAGAGACTGCGCGCTGTCTGCACTGCACTGCACTGCACTGCACTGCA 1118

Qy 341 ValGlyAlaPheGlnTYRGLNLeuHisLysValAsnIleSerAsnTYRSerLeuVal 360
Db 1119 GTTGTGCTTCAGATTTATTCAGTTACACAGGTAATATTTCCAACTCAGCTGCTG 1178

Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
Db 1179 ATGAGTCTACCGAGGAACTTCTTACCACTGCGGTGAATTCATCTGCAAGTGAAG 1238

Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuProLeuAsnIleThrGlnVal 400
Db 1239 GCCAGGCGCATTCATGAGAGAGTTATGAGCTGCTGCAAGCCCTCATATCATCAGTGC 1298

Qy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
Db 1299 CTGAGCCATGAGAGGAGCAGACATCACTTTTGATCCCAAGTGGAGTGGCGAGCC 1358

Qy 421 SerLeuLeuAspAspLeuTYRlysTYRPhPhePheHisHisSerHisGlyAspThrMet 440
Db 1359 AGCTTACTTATGACTTATACAGATATTTCTTCCATCATCTCCCAAGAGACCATG 1418

Qy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTYRAlaValIleSerTYR 460
Db 1419 ACTGTCAATGATCCAAAGCAGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTTAT 1478

Qy 461 ValValAlaAspMetGluMetLeuProArgSer 472
Db 1479 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1514

RESULT 3
AX006440
LOCUS AX006440 1860 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 5 from Patent WO0004157.
ACCESSION AX006440
VERSION AX006440.1 GI:9994575
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Neefs, J.M., Peeters, D.C. and Pangalos, M.
AUTHORS Cloning and characterization of novel mammalian peptidases
TITLE Patent: WO 0004157-A 5 27-JAN-2000;
JANSEN PHARMACEUTICA NV (BE); NEEFS JEAN MARC EDMOND FERNAND (BE);
PEETERS DANIELLE CELINE GEORGE (BE); PANGALOS MENILAS (GB)
FEATURES
source
1. 1860
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 518 a 420 c 451 g 471 t
ORIGIN
Alignment Scores:
Pred. No.: 1,196-175 length: 1860
Score: 2417.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-745-763-36 (1-472) x AX006440 (1-1860)
Qy 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCyseSerGly 20
Db 125 ATGAATTCCTTATCTTCGATTTTTCGATGTTTCACTTTTATCCCTGCTGCTGG 184
Qy 21 LysAlaIleCybLysAsnGlyIleSerLysArgThrPheGluGluIleLeuGluIle 40
Db 185 AAAGCTAATATGCAAAATGCGATCTCTTAAGAGACTTTTGAAGAAATGAATAA 244
Qy 41 AlaserCybGlyAspValAlaLysAlaIleLeuLeuAlaValTyrGlyLysAlaGln 60
Db 245 GCCAGCTGTGAGATGTTGCTTAAGCATCATCACTGCTTTATGTTAAAGCCAG 304
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db 305 AACAGATCTTAAGACGATGCGACTTCTGTTGATCTGTTGAGCCAGCTAGTGGC 364
Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnGlnAspGlyLeu 100
Db 365 TCCAGAACTTCAAGAAAGCCATCCAAATTATGTCCAAACCTGACAGATGGCTG 424
Qy 101 GlnLysValHisLeuGluProValArgIleProHisTyrPheGluArgGlyGluSerAla 120
Db 425 CAGAAAGTTTCACTGAGCGGAGTGAATACCCCTGAGAGAGGGAGAAAGATCACT 484
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 485 GTGAGCTGTGAGCCAAAGATTCATTAAGATGACCTCGGCTTGGCAGCAGCTGGG 544
Qy 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
Db 545 ACTCTCTCAGAAAGCATTAACAGCAAGATTCGTGGTGGTGAACCTCTTTCGATGA 604
Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrPheGlnProTyrIleAspTyr 180
Db 605 AGAAGGCTCTCAAGAGCAAGAGGAGATGTGTGTTTAAACCAACCTTACACTAC 664
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaLysValGlyAlaLeu 200

Db 665 TCAAGACGGTGCATATCCAGACGAGGGGCGGTGAGAGCTGCCAAGTTGGGGCTTTG 724
Qy 201 AlaserLeuIleAsgSerValAlaSerPheSerIleTyrSerProHisTnGlyIleGln 220
Db 725 GCATCTCTCATTCGATCCGTCGCTCTCTTCTTCATCTACAGTCTCTCACACAGGATTCAG 784
Qy 221 GlnTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 785 GAATACCAAGATGGCGTCCCAAGATTCCAAGACGCTGTTTACGGTGGAAAGTGCAGAA 844
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 845 ATGATGTCAAGAAAGGCTTCTCATGGATCAAAATTCATTCAGCTCAAAAGATGGGGCA 904
Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
Db 905 AAGACCTAACCAAGTACTGATCTTCAACACGTGTACAGAGATCATCTGGAGGCAATAT 964
Qy 281 ProGlnIleValIleValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
Db 965 CCAAGACAGGTGTACTGCTGTCAGTGACATCTGACAGCTGGGATGTTGGGCAAGGTGCC 1024
Qy 301 MetAspAspGlyGlyValAlaPheIleSerTyrGlnAlaLeuSerLeuIleLysAspLeu 320
Db 1025 ATGGATATGCGCGGTGAGGCTTTATATCATGGAGACCTCATCTTATTAAGATCTT 1084
Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrPheArgGluGlnGly 340
Db 1085 GGGCTGGCTCAAAAGAGACTCTGGCTGCTGTCTGAGCTGCAGAGAAACAAAGTGA 1144
Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
Db 1145 GTTGCTGCTTTCAGATTTATCATGATTACAGGTAATATTTCCAACTCAAGCTGCTG 1204
Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
Db 1205 ATGAGATCTGACGAGAGAACCTTCTTACCACTGGCTGCAATTCATCTGCAAGTGA 1264
Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuGlnProLeuAsnIleThrGlnVal 400
Db 1265 GCCAGGCGCATCATGAGAGGATTAAGCTGTGTGAGCCCTCAATATCATCTCAGGTC 1324
Qy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
Db 1325 CTGAGCATGAGAGAGGACAGACATCACTTTTGGATCCAGTGGAGTGGCCGAGGCC 1384
Qy 421 SerLeuLeuAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
Db 1385 AGTCTACTGATGACTTATACAGATATTTCTTCTTCATCATCATCCCAAGAGACACCATG 1444
Qy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValTyrPheAlaValIleSerTyr 460
Db 1445 ACTGTCAATGATCCAAAGACAGATGAATGTGCTGCTGTGTGGCTGTGTTTCTTAT 1504
Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db 1505 GTTCTTCAGACATGAGAAATGCTGCTCAGTGC 1540
RESULT 4
BC020689 1928 bp mRNA linear PRI 22-JAN-2002
LOCUS BC020689
DEFINITION Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418
IMAGE:4251802, mRNA, complete cds.
ACCESSION BC020689
VERSION BC020689.1 GI:18088383
KEYWORDS MGC.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Strausberg, R.
TITLE Direct Submission

JOURNAL

Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdickson@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

REMARK

COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNLT at: <http://image.llnl.gov>
Series: IRAL Plate: 35 Row: f Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706386.

FEATURES

source

Location/Qualifiers
1..1928
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:22418 IMAGE:4251802"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="NIH MGC_62"
/lab_host="DH10B"
/note="vector: pDNR-LIB"
167..1585
/codon_start=1
/product="plasma glutamate carboxypeptidase"
/protein_id="AAH20689.1"
/db_xref="GI:18088384"
/translation="MKFLIFAPFGVHLISCGKAIKNGISKRTPEIKETIASCG DVAKIINLAVYKGAQNSRYRLALINDTVPRSGSNLEKAIQIWNQIQGLEK VHAEPRIPIHMERGESAVMLEPRIHKIAIIGLSISITPEEGITAEVLTSTFDELQ RRAEARKGIVVNPQYINVSRTVQYRQGAVEAKAALSLRSVASFISYSPHTG IOEYODGVPIPTACITVEDAEEMMSRMAISHIKIYIQKMAKTYPDIDSRTNVAIT GSXPEOVVLVSGHLSMDVGQANDDGGAFFISWELSLIKDGLRKRLRLVLT AEEQGVGAFQYOLHKYNISNSIYNMSDAGTLPGLISGTSEKRAIMEEYMSLL QPUNITQYLSHGETDINFIQAGYCPGASLDDLDIKYFFHSHSGDTWTVDPKQWNY AAAYWAVVSYVVADEMLPRS"

CDS

BASE COUNT

554 a 435 c 466 g 473 t

ORIGIN

Alignment Scores:

Pred. No.: 1.24e-175 Length: 1928
Score: 2417.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-745-763-36 (1-472) x BC020689 (1-1928)

Qy 1 MetLysPheLeuLeuPheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20
Db 167 ATGAATATTCCTTATCTTCGATTTTCGGTGTTCACCTTTATCCCTGTGCTCTGG 226
Qy 21 LysAlaIleCyLysAsnGlyTLeSerLysArgThrPheGluGluIleLysGluGluIle 40
Db 227 AAAGCTATATCAAGAAATGGCATCTCTAAGAGCACTTTTGAAGAAATATAAGAAATA 286
Qy 41 AlaSerCyGlyAspValAlaLysAlaIleLeuLeuLeuAlaValTyrGlyValAlaGln 60
Db 287 GCCAGCTGTGGAGATGTTGCTAAAGCAATCAACCACTGAGCTTTAAGTGAAGCCAG 346
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80

Db 347 AACGATCTCATGAGCGATTCGACCTCTGTTGATACCTGTGGACCAAGTGAAGGCG 406
Qy 81 SerLysAsnLeuGlyLysAlaIleGlnIleMetCysTrpGlnLeuGlnGlnAspGlyLeu 100
Db 407 TCCAAAGACCTCAGAAAAGCCATCCAAATATGATCAAAAACCTGCAGCAAGATGGCTG 466
Qy 101 GlnLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluGlnSerAla 120
Db 467 GAGAAATTCACCTCGAGGCCAGTGAATACCCACCTCGAGAGAGAGAAATTCAGCT 526
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 527 GTGATCTGGAGCCAAAGATTCATTAAGATGACCTCTGGGTCTTGCAGACGATTCGG 586
Qy 141 ThrProGluGlyIleThrAlaGluValLeuValValHisSerPheAspGluLeuGln 160
Db 587 ACTCTCCAGAGGACATTACAGCAAGAGTTCTGTGTGACCTCTTTCATGAACTGCAG 646
Qy 161 ArgArgLaserGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
Db 647 AGAAGGCTCCAGAAAGCAAGCAAGGGAAGATTTGTTTATACCACTTACATCACTAC 706
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaLysValGlyValAlaLeu 200
Db 707 TCAAGACGGTGCAATACCGAAGCAAGGAGGCGGTGAGAGTCCAAAGTTGGGGCTTTG 766
Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db 767 GCATCTCTCATTCATCCATCCGGCTCTCTTCCATCAAGTCTTCCACAGGATTCAG 826
Qy 221 GlnTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 827 GAATACCAAGATGCGCGGCCCAAGATTCCAACGCTGTATTACGCTGGAAGAATGCAGAA 886
Qy 241 MetLysSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 887 ATGATGTCAGAAATGCGCTTCATGAGGATCAAAATTCATTCAGTCAAAAGATGGGGCA 946
Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
Db 947 AAGGCTTACCAAGATCTGATTCGTTCTTCAACCTGTACAGAGATCTGGGAGCAAAATAT 1006
Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
Db 1007 CCAAAAGAGTTGATCTGATGAGTGAACATCTGACAGCTGGAGTGTGGCAGAGGCTCC 1066
Qy 301 MetAspAspGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
Db 1067 ATGATGATGCGCGGTGAGCTTTATATCATGGGAACACTCTCATTTAAGATCTT 1126
Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnGlyGly 340
Db 1127 GGGCTGTGTCCAAAGAGAGACTCTGCGGTGTCTGTGACTGCAGAAACAGAGTGA 1186
Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValLeuIleSerAsnTyrSerLeuVal 360
Db 1187 GTTGCTGCTTCAGATATCACTTACAGTACACAGGATAATATTTCCAACTACAGTCTG 1246
Qy 361 MetGluSerAspAlaGlyTyrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
Db 1247 ATGAGTCTGACGAGAGAACCTTTTACCACTGGCTGCATTCATCGAGAGGAAAG 1306
Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db 1307 GCCAGGCGCATCAGGAGAGAGTATAGCTGCTGCAGGCCCTCAATATCACTCAGGTC 1366
Qy 401 LeuSerHisGlyGluGlyTyrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420
Db 1367 CTGAGCCATGAGAGGAGACAGACATCACTTTGATCCAAAGTGGAGTCCCTGGAGCC 1426
Qy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440

Db 1427 AGCTACTGATGACTTATACAGATTTCTTCCATCACTCCACGGAGACACCATG 1486
 Qy 441 ThrValMetAspProIySGlImeAsnValAlaAlaValTPAlaValSerTyr 460
 Db 1487 ACTGTCATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTGTTCTTAT 1546
 Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
 Db 1547 GTTGTGCAGACATGAGAAATGCTGCTAGTCC 1582
 RESULT 5
 BD105816
 LOCUS BD105816 1778 bp DNA linear PAT 18-SEP-2002
 DEFINITION Secretory protein.
 ACCESSION BD105816
 VERSION BD105816.1 GI:23200634
 KEYWORDS JP 2002502234-A/7.
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 1778)
 AUTHORS Jacobs, K., McCoy, J.M., Racie, L.A., Lavallie, E.R., Merberg, D. and Spaulding, V.
 TITLE Secretory protein
 JOURNAL Patent: JP 2002502234-A 7 22-JAN-2002;
 GENETICS INSTITUTE INC
 PN JP 2002502234-A/7
 PD 22-JAN-2002
 PF 16-APR-1997 JP 1997537384
 PR 18-APR-1996 US 08/634325, 13-JAN-1997 US 08/783520 PI
 KENNETH JACOBS, JOHN M MCCOY, LISA A RACIE, EDWARD R LAVALLIE, PI DAVID MERBERG,
 PI VIKKI SPAULDING
 PC C12N15/12, C07K14/47, A61K38/17
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers.
 FEATURES
 source 1..1778
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"
 BASE COUNT 514 a 386 c 422 g 456 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3 84e-175 Length: 1778
 Score: 2410.00 Matches: 471
 Percent Similarity: 99.79% Conservative: 0
 Best Local Similarity: 99.79% Mismatches: 1
 Query Match: 99.71% Indels: 0
 DB: 6 Gaps: 0
 US-09-745-763-36 (1-472) x BD105816 (1-1778)
 Qy 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20
 Db 17 ATGAAATTCCTTATCTTCGCAATTTTCGGTGTGTTCACTTTATCCCTGTGCTTGG 76
 Qy 21 LysAlaIleCyValAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40
 Db 77 AAAGCTATATGCAAGATGCGATCTCTAAGAGCACTTTTGAAGAAATAAAGAGAAATA 136
 Qy 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyValAsn 60
 Db 137 GCCACCTGTGAGATGTTGCTTAAAGCAATCACTAGCTGTTTATGTTAAAGCCAG 196
 Qy 61 AsnArgSerTyrGlyArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80
 Db 197 AACGATCTTATGACGATGCGACTTCTGTTGATCTGTTGGACCCAGACAGTAGTGGC 256
 Qy 81 SerLysAsnLeuGluLysAlaIleGluIleMetTyrGlnAsnLeuGlnAspGlyLeu 100

Db 257 TCCAAGAACCTAGAAAAAGCCATCCAAATTATGTACCAAAAAGCTGCAGCAAGATGGCTG 316
 Qy 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluSerAla 120
 Db 317 GAGAAAGTTCACTGGAGCCAGTGAATATCCCACTGGAGAGAGAGAAATCAAGCT 376
 Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
 Db 377 GTGATGCTGGAGCCCAAGAAATTCATTAAGATAGCCATCTCGGAGTCTGGCAGACGATTGG 436
 Qy 141 ThrProGluGlyIleThrAlaGluValLeuValHisSerPheAspGluLeuGln 160
 Db 437 ACTCTCCAGAGGCATTACAGCAGAAAGTCTGTGTGTGACCTCTTCATGATGAGCTCAG 496
 Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
 Db 497 AGAAGGCTCTCAGAGCAGAGAGGAGAAATGTTGTTTATACCACTTTCATCACTTAC 556
 Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyValLeu 200
 Db 557 TCAGAGACGGTGCATACCGAGACGAGGGGGGTGAGACTGTCAGAGTGGGGCTTTG 616
 Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleThrGlyIleGln 220
 Db 617 GCATCTCTCATTCGATCCGTGCTCTCTCTCCATCTACAGTCCCTCACAGAGTATTCAG 676
 Qy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 Db 677 GAATACAGAGATGCGCGCCCAAGATTCAGACGCTGATATAGGAGGAGAGACAGAA 736
 Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 Db 737 ATGATGTCAGAAATGCTTCATCAGGATCAAAATTCATTCATTCGTTAAAGATGGGCA 796
 Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
 Db 797 AAGGCTTACCCAGATGATGATCTCTTCACACCTGTACAGAGATCATCTGGAGAGCAATAT 856
 Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
 Db 857 CCAAGACAGGTGTACGTGCTGATGACATCTGCAGCTGGAGTGGAGTGGAGGAGTGGC 916
 Qy 301 MetAspAspGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
 Db 917 ATGATATGCGCGGTGAGGCTTTATATCATGGGAACACCTCATTTATTAAGATCTT 976
 Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnGly 340
 Db 977 GGGCTGCGTCCAAAGAGACTCTGCGCTGTGCTGTGAGCTGAGAGAGAAACAAGTGA 1036
 Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
 Db 1037 GTTGTGCTTTCAGATATATCATGTTACACAGGTAAATTTTCCATCTACAGTCTGGTG 1096
 Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
 Db 1097 ATGAGGCTAGACGAGAGAACTTTTACCACATGGCTGCATTCATCTGACAGAGAAAG 1156
 Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 Db 1157 GCCAGGCGCATCATGAGAGAGATTATAGCTGTGACGCCCTCATATACACTCAGATC 1216
 Qy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
 Db 1217 CTGAGCCATGAGAGAGAGACATCACTTATGATCCAGGTGAGTGGAGTGGAGGCC 1276
 Qy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisIleSerHisGlyAspThrMet 440
 Db 1277 AGTCTACTGATAGACTTATACAGATATTTCTTCTCATCACTCCACGAGACACCATG 1336
 Qy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValTPAlaValSerTyr 460
 Db 1337 ACTGTCAATGATCCAAAGCAGATGAATGTGCTGCTGCTGTTGGGCTGTGTTCTTAT 1396

Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db 1397 GTTGTTCAGACATGAGAAATGCTGCTAGTCTC 1432

RESULT 6
BD127520 1860 bp DNA linear PAT 18-SEP-2002
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD127520
ACCESSION BD127520
VERSION BD127520.1 GI:23222465
KEYWORDS JP 2002017375-A/2951.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1860)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.

REFERENCE
AUTHORS
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2951 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2951
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA

FEATURES
SOURCE
1. 1860
/organism="Homo sapiens"
/mol type="genomic DNA"
/db xref="taxon:9606"

BASE COUNT 515 a 423 c 451 g 471 t

ORIGIN

Alignment Scores:
Pred. No.: 1.66e-174 Length: 1860
Score: 2402.00 Matches: 469
Percent Similarity: 99.79% Conservative: 2
Best Local Similarity: 99.36% Mismatches: 1
Query Match: 99.38% Indels: 0
Gaps: 0

US-09-745-763-36 (1-472) x BD127520 (1-1860)

Qy 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20
Db 129 ATGAATTCCTTATCTTCGCAATTTTCGTCGTCGTCACCTTTATCCCTGCTGCTGG 188

Qy 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleGlyGluIle 40
Db 189 AAGCTATATGCAAGAAATGGCATCTCTTAAGGACTTTTGAAGAAATGAAGAAATTA 248

Qy 41 AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
Db 249 GCCAGCTGTGAGAGTGTCTTAAGCATCACTAGCTTTATGCTTAAGCCAG 308

Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db 309 AACAGATCTTATGACGATGGCTGCTGCTGTTACTGTTGAGACCAAGACTGAGTGGC 368

Qy 81 SerLysAsnLeuGluValAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100
Db 369 TCAGAAACCTTAAAGAAAGCATTCATATATGACAAAACCTGCAGCAAGATGGCTG 428

Qy 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluSerAla 120
Db 429 GAGAAAGTTCACCTGGAGCCAGTGAAGATACCCACCTGGGAGAGGAGAGATCAGCT 488

Qy 121 ValMetLeuGluProArgIleHisLysAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 489 GTGATGCTGAGCCAAAGAAATTCATTAAGATAGCCATCTGGGCTTGGCAGCAGATTGG 548

Qy 141 ThrProProGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160
Db 549 ACTCTCCAGAGAGCATTTACAGCAAGAGTTCTGGTGTGACCTTTTCAGATGACTCAG 608

Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
Db 609 AGAAGGCTCAGAGAGCAAGAGGAGGAGATGTTGTTTATTAACCACTTACATCACTAC 668

Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200
Db 669 TCAAGGACGGTGCATATCCAGACGAGGGCGGTGAGCTGCCAAGTTGGGGCTTTG 728

Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisTrnGlyIleGln 220
Db 729 GCATCTCTCATTCGATTCGGTGGCTCTTCTTCATCTCACTGCTTCCACAGGATTCAG 788

Qy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 789 GAATACAGAGATGCGTGGCCCAAGATTCCAACACCTGTATTACGGTGAAGATGCAAGAA 848

Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 849 ATGATGTCAGAAAGAGGCTTCTCATGAGATCAAAATTCATTCAGCTTAAGATGGGCA 908

Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
Db 909 AAGACCTACCCAGATCTGATCTCTTCAACACTGATACAGAGATCATCTGGAGCAAAAT 968

Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
Db 969 CCAGAAACAGGTGTGACTGTGACATCTGACATCTGGACGTGGATGTTGGGCAAGGTGCC 1028

Qy 301 MetAspAspGlyGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
Db 1029 ATGATGATGCGCGTGGAGCCTTTATATCATGGAGACACTCTCATTTAAGATCTT 1088

Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrTrpThrAlaGluGluGlnGly 340
Db 1089 GGGCTGCTCCAAAGAGAGACTCTCGGCTGTGCTCTGACTGCAGAAAGCAAGGTGGA 1148

Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
Db 1149 GTTGCTGCTCCAGATATATCACTTACACAGGTAATATTTCCACTACAGCTGCTGGT 1208

Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
Db 1209 ATGAGAGCTGACGAGAGAACTTCTTACCACTGGGCTGCAGATTCACCTGGCAGGAAG 1268

Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGluVal 400
Db 1269 GCCAGGGCCATCGTGAAGAGATTATAGCTCTGCGACCCCTCAATATATCTCAAGTCT 1328

Qy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
Db 1329 CTGAGCCATGAGAGAGAGAGACATCACTTTTGATTCACAGCTGGAGTGCCTGGAGCC 1388

Qy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
Db 1389 AGTCTACTTGATGATATATACAGATATTTCTTCTTCATCATCTCCACGAGAGACCATG 1448

Qy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTyrPheValValSerTyr 460

Db 1449 ACTGTCATGATCCAAAGACATGATGTCGTCGCTGTTGGCTGTTCTTAT 1508

Qy 461 valvalalaaapmetclucilmetleuproaargser 472

Db 1509 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1544

RESULT 7
AK075132

LOCUS

DEFINITION AK075132 1860 bp mRNA linear PRI 03-SEP-2002
Homo sapiens cDNA FLJ90651 fis, clone PLACE1004482, moderately similar to Rattus norvegicus hematopoietic lineage switch 2 related protein (H182-tp) mRNA.

ACCESSION AK075132
VERSION AK075132.1 GI:22761022
KEYWORDS oligo-capping; fis (full insert sequence).

SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS 1
1 Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T., Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S., Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kojima, S., Nagahara, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsubu, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Nimomiyu, K.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1860)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo; Laboratory of Genome Structure, Human Genome Center; cDNA 5'-3' end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).
Location/Qualifiers
1. 1860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004482"
/tissue_type="Placenta"
/clone_lib="PLACE1"
/note="Cloning vector: pME18FLJ3"
129. 1547
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC11423.1"
/db_xref="GI:22761023"
/translation="MKPLIPAPFGVHLISGKAIKCKISKRTPEIRERIASCS
DVAKIINLAYGKQNSRYERIALVVTYPRRLSGSKLKAIOIMYONQOQGLK
VLEPRVLPHEWGESAVMLEPRTHKIAIGLSSITPEGITAELVTSFDELQ
RRASBARGIIVNQPIYNSRTVQRTQGAWEAKVLAASLRSVASF1YSPHTG
IOEVODGPKIPIACTIVEDABEMSRMAHSGIKIVIQLMGAKTYPDIDSNTVAEIT
GSKYDEQVVLVSLHSDVSGAMDGGAFFISWEALFIKDLGARKRLRLVLT
AEQGVGAFOYVQLHKYINISYLSVMSDAGTFLPTGLQTTGSEKRAIYEEMSL
OPNITQVLSHGECDINFIQAGVPGASLDDLYKTFPHSHGDTWTWDPKQNV
AAAWAVSYVADMEMLPRS"

BASE COUNT 515 a 423 c 451 g 471 t

ORIGIN

Alignment Scores:
Pred. No.: 1.66e-174 Length: 1860
Score: 2402.00 Matches: 469
Percent Similarity: 99.79% Conservative: 2
Best local Similarity: 99.36% Mismatches: 1

Query Match: 99.38% Indels: 0
DB: 9 Gaps: 0

US-09-745-763-36 (1-472) x AK075132 (1-1860)

Qy 1 MetlyasphelenulephelaaphheglyyvalhiisleuenseleuCyseSergly 20

Db 129 ATGAAATTCCTTATCTTCGATTTTTCGGTGGTTCACCTTTATCTCTGCTGGG 188

Qy 21 LyvalalleCylysaanglylleserlysaargThrphelugluileysglugluile 40

Db 189 AAACCTTATGCAAGATGACATCTTAAAGAGACCTTTGAAGAAATGAAGAAATA 248

Qy 41 AlaserCyglyaspValaialysaiaileilasnleuvalaalyryglylysaalagln 60

Db 249 GCCAGCTGTGAGATGTTGCTTAAAGCATCATCACTTACTGTTTATGTTAAAGCCAG 308

Qy 61 AsnarSeryrygluarleuvalaleuValasprThryValglyProargleuSergly 80

Db 309 AACAGATCTATAGAGCATGGATGGACCTTCTGTTGATCTGTTGGACCAAGATGAGTGGC 368

Qy 81 SerlysaenleugluysaiaileglnilemetryrGlnleuenglninasrglyleu 100

Db 369 TCCAAAGAACCTTGAAGAAAGCCATCCAAATATGTAACCAAAACCTGCAAGATG33CTG 428

Qy 101 GlulyeValhlsleugluProvalArgileProhiestrpGluargglygluseral 120

Db 429 GAGAAATTCACCTGGAGCCAGTGAAGATACCCACCTGGAGAGGAGAGAGATCACTGCT 488

Qy 121 ValmetleugluProargylehlsylelaleilaleuglyleuSergSergly 140

Db 489 GTGATGCTGAGCCAAAGATTCATTAAGATAGCATCTGAGTCTGGACACAGCATGGG 548

Qy 141 ThrProglugluylethrralagluValleuValhysrPhesprleuSergln 160

Db 549 ACTCTCAAGAGCATTTAAGCAAGATCTGCTGATGATCTTTCAGATGAATGCTGAG 608

Qy 161 ArgarglaserGluAaargglylylvalValaValyrannglnProtyrileasrtyr 180

Db 609 AGAAGGCTCTCAAGAGCAAGAGGAGATGTTTATATACCACTTATCATCACTAC 668

Qy 181 SerargThryValglnyryargThrglnlylaValagluAaalyValaglualeu 200

Db 669 TCAAGAGCGGTGCAATCCAGACGAGGCGGTGGAACCTGCAAGGTGGGCTTTG 728

Qy 201 AlaserleuleargserValaaserPheserileyrserProhiesthrglyleGln 220

Db 729 GCATCTCTCATTCGATCCGTCCTCTCTCTCATTCACGTCCTCACAGGATTCAG 788

Qy 221 GlutyrglnasrglyValaProlyslleProThralaCysileThryValgluaspaaglu 240

Db 789 GAATACAGATGGCGGCCCAAGATTCACAGCCGCTTATTCGGTGAAGATGCAGAA 848

Qy 241 MetwetseryargmetalasernhsaglyllylelylvallelgleuleuSmetGlyla 260

Db 849 ATGATGTCAGAAATGCTTCTCATGGGATCAAAATGTTCATTCAGCTTAAAGATGGGCGCA 908

Qy 261 LyethrtyrProasprthrapserPhesanthValaaleuilethrglyserlytyr 280

Db 909 AAGGCTTACCAAGATTCATTCCTTCACACGTCGACAGATCATCTGGAGCAATATAT 968

Qy 281 ProgluglnValaalleuValaSerGlyhlsleuAspserThraspValglynglyla 300

Db 969 CCAGAACAGTTGTACTGCTGACATGAGACATCTGACAGCTGGAGATGTGGGAGGCTGCC 1028

Qy 301 MetaspaapglglylylaalaphelleSertrpGlualeuSeriuleyaspLeu 320

Db 1029 ATGATGATGCGCGGTGAGCCTTATATATCATGGGAAGCATCTCACTTATTAAGATCTT 1088

Qy 321 GlyleuargProlysaargThrleuargleuValleuThrThralagluuglnnglyly 340

Db 1089 GGGCTGCTGTCAAAGAGACCTGCGGCTGCTGCTGACCTGACAGCAAGAAAGAGTGA 1148

	(NMAIDbase) : similar to sequences with Genbank Accession Numbers R18560 and G07404; prostate-specific membrane antigen (PSMA)"
	/codon_start=1
	/product="blood plasma glutamate carboxypeptidase precursor"
	/protein_id="FADJ1418.1"
	/db_xref="GI:4877698"
	/translation="MKFLIPAFPGVHLISLCSGKAI CKNGISKRTPEEIKKEIASCGC/DVAKIINILAVYKAQNRSEYERIALVDVTGPRLSGSKNLKAI QIWMYQMLQDGERKVHLEPVAIPMERGEESAVMLEPRIRHTAIIALIGSSSIGTPEGITAAVLVTSDELR/RRASEARGLIVYNOPINYSRTVYRTOGAVALAKAGALASILRSVASISYPHGIGORYODSVPIPTACTIVEDAEEMMSMASGISIKIVIOLKMGAKEYPTDTSFNVAETTGSKRYPEGVAVSGLHSMDVYGCMNDGGCAFISWEALSILKDLGRPKTTLVLT/AEEQGVGARFOYIQLHKVNISNYSLVNESDAGFTLPGLQPTOSEKARAIMEEWSLL/QPLNIQVLSHGEGTDINFMIQAQVPASLDLDLYKFPHSHGDMTWDPBRWL/LIFGLFLMLLOTWKCCCLGPNRSKSETSCFWPGTILQLMWTPLHITISSNSSKK/HNSICFLILSFILSRFSRKRNSPLRPPT"
sig_peptide	25..96
	/gene="PGCP"
mat_peptide	157..1647
	/gene="PGCP"
	/product="blood plasma glutamate carboxypeptidase"
misc_feature	205..207
	/gene="PGCP"
	/note="putative; glycosylation site"
misc_feature	559..561
	/gene="PGCP"
	/note="putative; glycosylation site"
misc_feature	1081..1083
	/gene="PGCP"
	/note="putative; glycosylation site"
misc_feature	1090..1092
	/gene="PGCP"
	/note="putative; glycosylation site"
misc_feature	1210..1212
	/gene="PGCP"
	/note="putative; glycosylation site"
BASE COUNT	522 a 391 c 422 g 457 t 2 others
ORIGIN	
Alignment Scores:	
Pred. No.:	3.79e-173 Length: 1794
Score:	2384.00 Matches: 470
Percent Similarity:	99.79% Conservative: 1
Best local Similarity:	99.58% Mismatches: 1
Query Match:	98.63% Indels: 1
DB:	9 Gaps: 0
US-09-745-763-36 (1-472) x AF119386 (1-1794)	
Oy	1 MetlysPhenleuilePheAlaPhePheGjVgLyValHisLeuLeuSerLeuCysSerCjly 20
Dbb	25 ATGAAATTCCTTAATCTTCGCATTTTTGGTAGGTGTTCACTTTATCCCTGTGCTCGGG 84
Oy	21 LysAlaileCyBjlySaangilyileSerlysaArgThrPhegiugiuiletysgjugnuile 40
Db	85 AAACGTATATGCAAGAATGGCATCTCTTAAGAGCACTTTTGAAGAAATAAAAAGAAATA 144
Oy	41 AlAserCYegijYaapValAlalyAlailelleasnLeuAlaValIlyrgjlyValAgin 60
Db	145 GCCAGCTGTGAGATGTTGCTTAAAGCAATCATCAACTGACTGTTTATGCTTAAAGCCGAG 204
Oy	61 AsnaIrgserTyrgIuaTgyleuAlaleuleuValAspThrValIgIyrOargLeuSerCjly 80
Db	205 AACGATTCCTTAATGAGCCATTGCGACTTCTGTGTATACTGTTGAACCCAGACTAGTGGC 264
Oy	81 SerlysaSnleugiulySaAlaileginllewertyrgInbnleugiinginaSpGjlyeu 100
Db	265 TCCAAGAACCTAGAAAAAGCATCATCAATATTATGTAACAATAAACCCTGACAAAGATGGGCTG 324
Oy	101 GIUuySaValHisLeuGiUpProvalArgilleProhiStrPgIuaTggjlygiugiuSerAla 120

```

Db      325 GAGAAAGTTCACTTGAGCCAGTGAATACCCCACTGGAGAGGAGAAATCAGCT 384
Oy      121 ValMetLeuGluProArgIleHisIleValIleIleLeuGlySerSerIleGly 140
Db      385 GTGAGCTGGAGCCAGAAATTCATAGATAGCCATCTGGGCTTGGAGCAGCATTTGG 444
Oy      141 ThrProGluGlyIleThrAlaGluValIleValIleThrSerPheAspGluLeuGln 160
Db      445 ACTCTCCAGAAAGCATTTACAGCAGAGTTCTGGTGGTGAAGCTCTTGGATGAATGAG 504
Oy      161 ArgGalaSerGluAlaArgGlyIleValIleValIleThrGlnProTyrIleAsnTyr 180
Db      505 AGAAGGGCTCAGAAAGCAAGAGGAAAGATTGTTGTTTATACCAACTTACCTACATC 564
Oy      181 SerArgThrValGlnTyrArgThrGlnGlyAlaValIleAlaIleValIleValIle 200
Db      565 TCAGGAGGCTGCAATACGAGAGGAGGGGGGGAGGAGGCGCCAGAGGGGGGCTTGG 624
Oy      201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db      625 GCATCTCTCATTCGATCCGGTGGCTCTCTTCTCATCTACAGTCCCTCAGACAGATTCAG 684
Oy      221 GluTyrGlnAspGlyValProIleIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      685 GAATACCAAGATGGCGGCCCAAGATTCACACGCTGATTAACGTTGAAGATGCAGAA 744
Oy      241 MetMetSerArgMetAlaSerHisGlyIleIleValIleGlnLeuIleMetGlyVal 260
Db      745 ATGATGTCAGAGATGCTTCTCATGAGTCAAAATTCATTCACCTAAAGTGGGGCA 804
Oy      261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280
Db      805 AAGACCTACCCAGATGACTGATCTCTTCAACATGTAAGCAGAGATCAGTGGAGCAAAAT 864
Oy      281 ProGluGlnValIleValIleSerGlyHisLeuAspSerTTPAspValIleGlnGlyAla 300
Db      865 CCAGAACAGTTGTCTGCTGAGTCTGAGACATCTGAGACACTGGAGTGGGAGCGGTGCC 924
Oy      301 MetAspAspGlyIleValIleAlaPheIleSerTTPGluAlaLeuSerLeuIleIleAspLeu 320
Db      925 ATGGATGATGGGGGAGGAGCTTTATATCATGGAGACACTCTCACTTAATTAAGATCTT 984
Oy      321 GlyLeuAspProIleArgThrLeuArgLeuValIleTTPThrAlaGluGlnGlyIle 340
Db      985 GGGCGGCTCCAAAGAGACTCTGGGCTGGTCTGAGACGCGAGAGAAAGAGAGTGA 1044
Oy      341 ValIleAlaPheGlnTyrTyrGlnLeuHisIleValIleSerAsnTyrSerLeuVal 360
Db      1045 GTTGGTGGCTTCCAGTATTCATGATTAACAAGTAAATATTTCCAACTACAGTCTGGTG 1104
Oy      361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluIle 380
Db      1105 ATGGAGTCTGAGGAGGAACTTCTTACCACTGGGCTGCAATTACCTGGCAGTGAAG 1164
Oy      381 AlaArgAlaIleMetGluGluValIleMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1165 GCCAGGCGCATCATGAGAGAGGTTATGACCTGCTGCAAGCCCTCAAAATTCATCAGATC 1224
Oy      401 LeuSerHisGlyGlnGlyThrAspIleAsnPheTPIleGlnAlaGlyValProGlyAla 420
Db      1225 CTGAGCCCATGAGAGGAGCAGACATCAACTTTGGATCCAACTGAGAGTGGCTGAGACC 1284
Oy      421 SerLeuLeuAspAspLeuTyrTyrPhePhePheHisIleSerHisGlyAspThrMet 440
Db      1285 AGTCTACTTGATGACTTATACAGATTTCTTCTTCCATCACTCCACAGGAGACCAATG 1344
Oy      441 ThrValMetAspProIleGlnMetAsnValAlaAlaIleValIleValIleValIle 460
Db      1345 ACTGTCATGATCCAG--CAGATGATGATGCTGCTGCTGTTGGGCTGTTCTTAT 1403
Oy      461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db      1404 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1439

```

```

RESULT 9
AF107835
LOCUS      AF107835      1596 bp      mRNA      linear      ROD 12-JUL-1999
DEFINITION Mus musculus clone LCH-M1 aminopeptidase mRNA, complete cds.
ACCESSION AF107835
VERSION    AF107835.1 GI:5442031
KEYWORDS
SOURCE
ORANISM    Mus musculus (house mouse)
REFERENCE
  1. (bases 1 to 1596)
  Liu, C.H., Lin, B.Y. and Chang, L.Y.
  Cloning of the mouse aminopeptidase gene
  Unpublished
  2. (bases 1 to 1596)
  Liu, C.H., Lin, B.Y. and Chang, L.Y.
  Direct Submision
  Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
  Sinica, RM 433, 128 Yen-Chiun-Yuan Road SEC 2, Taipei 11529, Taiwan
FEATURES
  source
    1..1596
      /organism="Mus musculus"
      /mol_type="mRNA"
      /db_xref="taxon:10090"
      /clone="LCH-M1"
      /tissue_type="liver"
      /note="secretory protein"
      /codon_start=1
      /product="aminopeptidase"
      /protein_id="P463215.1"
      /db_xref="GI:5442032"
      /translation="MRSPLFLFVHLALDGSQAVFKNVQSORTFREIKEIANYEDV
      /AKA1INLAVVGYQNSYERLGLVDVGRLEGSKNLEVAIIMQNLQDDLENAH
      LEQRIIPMERGESAMLEPRILKMAILLGSSIGTPGIIAEVLVVASFPELQRP
      ASEARGKILIVNOPYTGVEKTVQRYVQGAVEAKVAVASLIQSASFISPHGTIO
      KYDDGVKRIPTACTIVEDAEMSRMSRGKVIYHLEMGAKTYPDTSFMTVAEITTS
      MYREEVVLVSGHLDSDVDYGGALDDGGATISBAUSLVVDICLRPRTIRLVMTRE
      EQGIGASQYIEHLKANIISKYSLVMEADSGTFLPTGQIFGSDAKARIMKEVNNLLQP
      LNVTKVFSNDEGTDINFWIAGVPGASLRDILYKYPFFHSHSDPTMTVMDPKQNNVAA
      AVAAVAVVAVVADMDMLPRS"
BASE COUNT      415 a      357 c      421 g      403 t
ORIGIN
Alignment Scores:
  Pred. No.:      5.14e-154      Length:      1596
  Score:          2133.00      Matches:      414
  Percent Similarity: 93.64%      Conservative: 28
  Best Local Similarity: 87.71%      Mismatches: 28
  Query Match:    88.25%      Indels:      2
  DB:             10      Gaps:      1
US-09-745-763-36 (1-472) x AF107835 (1-1596)
Oy      1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisIleLeuSerLeuCysSerGly 20
Db      184 ATGAGGTCCTCTTCTTCTCTGTTCAAT-----GTACCTTTTAAAGCTTGGGCTGGT 237
Oy      21 LysAlaIleCysLysAsnGlyIleSerTyrArgThrPheGluGluIleLysGluGluIle 40
Db      238 AAAGCTGATTCAGAGATGAGTGTCTTCACAGCAATTTGAGAGAAATAAAGAGAAATA 297
Oy      41 AlaSerCysGlyAspValAlaIleValAlaIleIleAsnLeuAlaValIleGlyLysAlaGln 60
Db      298 GCCAATATGAAAGATGTGGCAAAAGCAATATCAACCTTCTGTTATGTGTAATACAG 357
Oy      61 AsnArgSerTyrGluGluGluAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db      358 AACCGGTCCTATGAGCGCTTGGAGCTTCTAGTTGATACCTGTGAGACCCAGACTGATGGC 417
Oy      81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100

```

```

|||||
418 TCTAAGACCTAGAGAAAGCTATTCAATCATGTATCCAAAACCTGCAGCAAGATGGCTTA 477
Qy 101 GlnuylValHstleuGluProValArgLleProHstPgiuArgLylGluGluSerAla 126
Db 478 GAAATATTCACTCGAGACAGAGTCAAGATACCCACCTGGAGAGGAGGAAATCTTGCA 537
Qy 121 ValMetLeuGluProArgLleHstleuAlleAlleuGluLeuGlySerSerLleGly 140
Db 538 GTGATGCTTAGCTCCATTCACAAAGATGCTTTCAGTCTTGGCAGACGATTTGGG 597
Qy 141 ThrProGluGluGlyIleThrAlaGluValleuValThrSerPheAspGluLeuGln 160
Db 598 ACTCCCGCAGAGGACATCACAGCAAGATGCTGGTGGAGCTCTTTGATGAATTCAA 657
Qy 161 ArgArgAlaSerGluAlaArgGlyLylleAlleValIleArgGlnProGlyIleAsnTyr 180
Db 658 AGAAGACATCAGAGCAAGAGGAGAGATCTTTATATACAGCCTTACACTGGCTAT 717
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaIleAlaIleValAlaLeu 200
Db 718 GAGAGACTGTGCATGACCGGTGCAAGGAGCTGTGAGCTGCCAAGGTGGAGCTGTG 777
Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerLleTyrSerProHstHstGlyIleGln 220
Db 778 GCATCCCTCATCAGATCAGTACGTCTTCCATCTACAGTCTCTCACACAGTATTCAG 837
Qy 221 GlnuylThrAspGlyValProLylleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 838 AAAATATCAGAGTGTGTGCCAATATCCACAGCTGTATCACAGATGAGAGTGCAGAA 897
Qy 241 MetMetSerArgMetAlaSerHstleuGlyIleLylleValleAlleGlnLeuMetGlyAla 260
Db 898 ATGATGCTCGAATGCTTCTCGTGGAAACAAATTCATTCATCTGGAATGGAGCA 957
Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleHstGlySerLylTyr 280
Db 958 AAGACATATCCAGATACATGATCTCTTCACTGTTGAGAGATCATCTGGAGCATGTAT 1017
Qy 281 ProGluGlnValIleValleuValSerGlyHstleuAspSerTyrPaspValGlyGlnGlyAla 300
Db 1018 CCAAGAGAGATTTGCTCGTGCAGTGACATTTGACAGCTGGAGTGTGGCAGAGTGCA 1077
Qy 301 MetAspAspGlyGlyValAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
Db 1078 CTGATGATGTGTGTGGAGCTTCAATCATGAGGAGACATCTCACTGTATTAAGATCTT 1137
Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValleuTyrThrAlaGluGluGlnGlyGly 340
Db 1138 GGGCTGGGTCCAAAGAGGACTCTGCCCTGTGCTGTGACTGCAGAGAAAGAGAGGA 1137
Qy 341 ValAlaIlePheGlnTyrTyrGlnLeuHstleuValleuIleSerAsnTyrSerLeuVal 360
Db 1198 ATTGTGTCTCCCGATTTATGACTCATAGCCAAATATTTCCAGTATCAGCTGTGGTG 1257
Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380
Db 1258 ATGAGAGTGCATCAGGAACCTTCTTACCACTGCAGCTGCACCGGAGAGGACAG 1317
Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db 1318 GCCAGGGCTATCAGAGGAGAGTCAATGATCTTCTGCACACCCCTCAATGATCACCAGGTC 1377
Qy 401 LeuSerHstleuGlyThrAspIleAsnPheThrPileGlnAlaGlyValProGlyAla 420
Db 1378 TTTGTATGAGAGAGAACTGATTTATTAATCTTGATCCAAAGCTGGAGTACCTGGAGCC 1437
Qy 421 SerLeuLeuAspAspLeuTyrLylleThrPhePheHstleuSerHstleuGlyAspThrMet 440
Db 1438 AGCTGCGCTGATGATGTACAGATTTCTTTTTCATCATCTCCACGAGGAGACCAATG 1497
Qy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValIleValIleValSerTyr 460

```

```

Db 1498 ACTGTCAATGATCCAAAGCAGATGAATGTTGCTGCTGCTGTATGAGGCTGTGTGCTTAT 1557
Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db 1558 GTTGTTCAGACATGATGAATGCTGCTTACGATTC 1593

RESULT 10
BC037067
LOCUS
DEFINITION
MUS musculus plasma glutamate carboxypeptidase, mRNA (cDNA clone
WGC:46857 IMAGE:4989536), complete cds.
VERSION
BC037067
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)

REFERENCE
AUTHORS
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buettow K.H., Schemper C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Schaefer T.E., Brownstein M.J., Udutin T.B., Toshiyuki S.,
Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Muliyil S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettleman M., Madan A., Young A.C., Shvachenko Y.,
Sanchez A., Whiting M., Madan A., Young A.C., Shvachenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalski U., Smalius D.E.,
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.,
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2388257
12477932
2 (bases 1 to 1796)
Strausberg, R.
Direct Submission
Submitted (23-AUG-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NHI-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@hgrl.nih.gov
Akhter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaesi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastriani, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stancik, S., Thomas, P.J., Touchman, J.W.,
Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAX Plate: 81 Row: 9 Column: 18
This clone was selected for full length sequencing because it

```

passed the following selection criteria: matched mRNA gi: 9055233.

FEATURES
source location/Qualifiers

1..1796
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:46857 IMAGE:4989536"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_fld="NCI_CGAP_Co24"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1796
/gene="Pcgp-pending"
/note="synonym: Hfs2, 1190003p2.rik"
/db_xref="locusID:54381"
/db_xref="MGI:1889205"
63..1475
/product="plasma glucose carboxypeptidase"
/protein_id="AAH37067.1"
/db_xref="GI:22477498"
/db_xref="LocusID:54381"
/translation="MRSLEFLFVHLALGSGKAVKNGVSQRTFERIKEIANEDV
AKAIIINLAVYKQKRSYERLGLVDTPGLSGSKNEKAIQIWMYOLQODLENVH
LEOARIPIHMERGESAVMLBPIHKMALLIGSSIGTPGGITAEVLSVSPHTGLO
ASEARGLIIVNOPPYGYEKTYOVQCAVEAKGAVASLIQSASISPHIGIO
KYODGPKIPACTIVEDENMRASGKVIHLEKAGKTPEDTSFNTVAETGS
MYPEBIVLHGLSDWDVQCALDGGAFISWEALSLVKDLRPKKTLVLMTAP
EOGIGASQVYELHAKNISKYSLVMEADSGTFLPGILOFTGSDKARAIKEMVLLQF
LNTVTFNNGESTDINFMIQAGVPGASLRDLYKFFPHSHSGDTMWDPKQNVAA
AVMAVAVVAVMDMEMLPRS"

CDS

BASE COUNT 526 a 372 c 418 g 480 t

ORIGIN

Alignment Scores:

Pred. No.: 5,966-154 Length: 1796
Score: 2133.00 Matches: 414
Percent Similarity: 93.64% Conservative: 28
Best Local Similarity: 87.71% Mismatches: 28
Query Match: 88.25% Indels: 2
Gaps: 1

US-09-745-763-36 (1-472) x BC037067 (1-1796)

QY 1 MetLysPheLeuIlePheAlaPheheGlyGlyValHisLeuLeuSerLeuCySerGly 20
DB 63 ATGAGCTGCTCTTCTCTCTGTCATT-----GTTCACTTTTGGCTTGGCTCTGCT 116
QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleGlyGluIle 40
DB 117 AAAGCTGATTCTCAAGAAAGGTGTTCTCAGCAACCTTTCGAAATTAAGAAAGAAATA 176
QY 41 AlaSerCySglYAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
DB 177 GCCAATCTAAGAGATGTTGCTTAACCAATTATCACTTCTTAAAGTAAATACAG 236
QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
DB 237 AACCGTCTAAGACGTTTGGGACTTCACTTGAATGTAAGTGGAGCCAGACTGAGTGC 296
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100
DB 297 TCTAAGAACTTAAGAAAGCTATTCAATCAATGATGACAAACCTCAGAGATGGCTCA 356
QY 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerAla 120
DB 357 GAAATGTTCACTCGAGAGAGCTCAGATACCCCACTGGAGAGAGGAGAAATCTGCA 416
QY 121 ValMetLeuGluProArgGlyLeHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
DB 417 GTGATGCTTGAAGCTCGAATTCACAAGATGCTATTCTAGGCTTGGAGAGCAGCATGGG 476

QY 141 ThrProGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160
DB 477 ACTCCCCAGAGGACATCAGCAGAAAGCTGCTGCTGTTGATGAACTTCA 536
QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
DB 537 AGAAGACATCAGACAGAGAGGAGGAGATCTGTTATTAACCGACTTACACTGCTAT 596
QY 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaLysValGlyValLeu 200
DB 597 GAGAAAGACTGCGTACCGGTCGACAGAGCTGTGAAGCTGCAAGGTGGAGCTGTG 656
QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
DB 657 GCATCCCTCATCCATCAGATAGTACTCTCTTTCATCTACAGCTCTCACAGATTCAG 716
QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
DB 717 AATATCAAGATGCTGTCGCCAAGATTCCAACAGCTGTATCAGATGAAGATGACAGAA 776
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyVal 260
DB 777 ATGATGCTCGAATGCTCTCTCTGGAACAAATTTTCATCTCTGAAATGGAGCA 836
QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
DB 837 AAGACCTATCCAGATCTGATCTCTTCAATCTCTTCCAGATCATCTGGAGAGATGAT 896
QY 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyVal 300
DB 897 CCAAGAGAGTGTCCGTGTCAGAGCAATTTGACAGCTGGAGTGTGGGAGGAGTGA 956
QY 301 MetAspAspGlyGlyValAlaPheIleSerThrGluAlaLeuSerLeuIleLysAspLeu 320
DB 957 CTGATGATGTGTGGAGCTTCTCATATCATGGAACACTCTCACTGTTAAAGATCTT 1016
QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnGlyGly 340
DB 1017 GGGCTGCTCCAAAGAGAGACTCTGCCCTGCTGCTGAGCTGCAAGAAAGAGAGGA 1076
QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
DB 1077 ATTGCTCCCTCCAGATTTATGAGCTATAGAGCAATATTTTCCAGATCAGCTTGCTG 1136
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
DB 1137 ATGGAAGCTGACTGAGAACCTTCTTACCACTGGAGCTGACGTTCACTGGAGTGAAG 1196
QY 381 AlaArgAlaIleMetGluGluValIleMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
DB 1197 GCCAGGCTATCATGAAGAGAGTATGATCTTCTGCAACCTCCATATGTCACCAAGTCT 1256
QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyVal 420
DB 1257 TTTGATATGAGAGAGAGACTGATATTAATCTTCTGATTCACAGCTGAGTACCTGGAGCC 1316
QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePheHeHisSerHisGlyAspThrMet 440
DB 1317 AGCTGCGTATGACTTGTCAAGATATTTCTTTTTCATCATATTCACGAGAGACCATG 1376
QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTyrAlaValSerTyr 460
DB 1377 ACTGTCATGATCCAAAGCAGATTAATGTTCTGCTCTGATGAGCTGTTGTGCTTAT 1436
QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
DB 1437 GTTGTTCAGACATGATGAATGCTGCCAGGTCC 1472

RESULT 11
AF131077 1726 bp mRNA linear ROD 03-JAN-2003
DEFINITION Rattus norvegicus liver annexin-like protein (LAL) mRNA, complete cde.

ACCESSION AP131077 GI:7108712
VERSION AF131077.1
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1726)
AUTHORS Della Fazio M.A., Piobbico D., Bartoli D., Castellani M., Brancorsini S., Viola Magni M. and Servillo G.
TITLE la1-1, a differentially expressed novel gene during proliferation in liver regeneration and in hepatoma cells
JOURNAL Genes Cells 7 (11), 1183-1190 (2002)
MEDLINE 2278398
PUBMED 12390252
REFERENCE 2 (bases 1 to 1726)
AUTHORS Servillo G., Della Fazio M.A., Piobbico D., Bartoli D., Castellani M., Brancorsini S. and Viola Magni M.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1999) Institute of General Pathology, University of Perugia, Policlinico Monteluce, Perugia 06100, Italy

FEATURES
source location/qualifiers
1..1726
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..1726
/gene="LAL"
63..1481
/gene="LAL"
/note="Involved in liver regeneration; contains unique annexin domain at the N-terminus"
/codon_start=1
/product="liver annexin-like protein"
/protein_id="AA036518.1"
/db_xref="GI:7108713"

gene
CDS
474 a 385 c 428 g 439 t

BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 9,49e-153 Length: 1726
Score: 2117.00 Matches: 413
Percent Similarity: 92.58% Conservative: 24
Best Local Similarity: 87.50% Mismatches: 35
Query Match: 87.59% Indels: 0
Gaps: 0

US-09-745-763-36 (1-472) x AF131077 (1-1726)

Qy 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCyseSergly 20
Db 63 ATGAGGTTCTTTCTTCTCTGTTGCTGTTGTTCACTTTCTCTCTGAGGCTTGA 122

Qy 21 LysAlaIleCySlyVaSnglyYlIseSerySargThrPheGluGluIleLysGluIle 40
Db 123 AAACCTATATCAAGAGTGGTTCCTGCGCAATTTCAAGAAATAAAGAAAGAAATA 182

Qy 41 AlaseGysGlyAspValAlaIleValIleLeuLeuAlaValIleGlyLysAlaGln 60
Db 183 GCCAATATGAGATGTTGCTTAAGCAATTTCACTTCTGTTTATGGAATAATACAG 242

Qy 61 AsnArgSeryTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSergly 80

Db 243 AACCGTGTATGAGCGTTTGGACCTTCTAGTATATCTTGGACCCAGACTGATGCC 302
Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
Db 303 TCTAAGAACTTAGAGAAAGCTTATCCAAATATGTTACCAAACTGCAACAGATGGCTG 362

Qy 101 GluLysValHisLeuGluProValArgIleProHisIleTyrGluArgGlyGluSera 120
Db 363 GAAACCTCACCTGAGCAGAGTCAAGATATCTTCACTGGAGGAGGCGAAGATCTGCA 422

Qy 121 ValMetLeuGluProAlaGlieHisLysIleAlaIleLeuGlyLeuGlySergly 140
Db 423 GTGATGTGTGGCTCCAAATTCACAACTGTATTTAGCCCTTGGCGGAGCATGGG 482

Qy 141 ThrProGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160
Db 483 ACTCTCTTAAGGTATCACAGCAAGATCTCGTGTGGCTTTTGTGAACTTCA 542

Qy 161 ArgArgLaseGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
Db 543 AGAAGGCATACAGAGCGAAGGGAATTTGTTTATACACGCTTACACTGACTAT 602

Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValAlaIleValAlaValLeu 200
Db 603 GGGAAACTGTGCAGTACCGGAGCGGAGCTGTGGAAGCTGCCAAGTGGGGCCGTG 662

Qy 201 AlaseGluLeuAspSeraValAlasePheSeryIleTyrSerProHisIleGln 220
Db 663 GCATCCCTCATCGATGATGAGCTCTTTTTCATCTACAGTCTTCAACAGGTCATCAA 722

Qy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAla 240
Db 723 GGATATCAAGATGTGTGCCCAAGATTCCAACAGCCGTGTACAAATAGAAAGCGA 782

Qy 241 MetLysSerArgMetAlaseHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 783 ATGATGTCTGCAACGCTTCTCGTGGGCAAAATTCATTCATCGAAATATGGAGCA 842

Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySeryTyr 280
Db 843 AAGACCTATCCAGATACAGATCTCTTCAACAGCTGTCCAGAGATCAGTGGAGCA 902

Qy 281 ProGluGlnValIleLeuValSerglyHisLeuAspSeryTyrAspValGlyGlnVal 300
Db 903 CCAAGGAATGTCTCGTCAAGTCACTGCAACGCTGGAGCGTGGCAGAGTCCA 962

Qy 301 MetAspAspGlyGlyAlaPheIleSeryTyrGlnAlaLeuSeryLeuIleLysAspLeu 320
Db 963 CTGATGATGCGCGGTGAGCCTTCAATCATGGGAACACTTCATCTTTAAAGATCTT 1022

Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrPheArgGluGlnGly 340
Db 1023 GGGCTGCTCCAAAGAGACTCTGGCTGTGTCTGACCGGAGAAACAAAGAGG 1082

Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValLeuIleSeraAsnTyrSeryVal 360
Db 1083 GTTGTCCTCCCAAGTATATGACCTATAGGCAAAATTTTCCAGTACAGTTGGTG 1142

Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySeryLys 380
Db 1143 ATGAGGCTGATCAGGAACCTTCTTACCACTGGCTGACAGTTCACCGGACAGTCAAG 1202

Qy 381 AlaArgAlaIleMetGluGluValIleMetSeryLeuGlnProLeuAsnIleThrGlnVal 400
Db 1203 GCCAGGCTATCAAGAGGAAGTATAGTCTCTGGAACCCCTCAATATCACCAAGGTC 1262

Qy 401 LeuSeryHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420
Db 1263 TTTATATATGAGAGAGAACTGACATTAATCTTCGATTCACAGTGAAGTCCCTGGAGCC 1322

Qy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSeryHisGlyAspThrMet 440
Db 1323 AGTCTGAGAGATGACTGTACAGATATTTCTTTCCATCATCTCCATGAGACACCATG 1382

QY 441 ThrValMetAspProLySGlnMetAsnValAlaAlaValTPrAlaValIserTyr 460
 DB 1383 ACTGCATGATGCCAAAGCATGAAATGTCTGCTGCTGTTGGGCTGTCTTCCCTTAC 1442
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
 DB 1443 GTTGTGCAGACATGAGGAAATGCTGCCAGCTCC 1478
 RESULT 12
 AX400829
 LOCUS AX400829 1778 bp DNA linear PAT 06-JUN-2002
 DEFINITION Sequence 505 from Patent WO0210453.
 ACCESSION AX400829
 VERSION AX400829.1 GI:21337009
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1
 AUTHORS Mendrick,D., Porter,M.W., Johnson,K.R., Castle,A.L. and
 Elashoff,M.R.
 TITLE Molecular toxicology modeling
 JOURNAL Patent: WO 0210453-A 505 07-FEB-2002;
 Gene Logic, Inc. (US)
 FEATURES
 source
 1..1778
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /note="EMBL/GenBank Accession No. AF097723"
 BASE COUNT 488 a 391 c 442 g 457 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,186-152 Length: 1,778
 Score: 2116.00 Matches: 413
 Percent Similarity: 92.58% Conservative: 24
 Best Local Similarity: 87.50% Mismatches: 35
 Query Match: 87.55% Indels: 0
 Gaps: 0
 US-09-745-763-36 (1-472) x AX400829 (1-1778)
 QY 1 MetLysPheLeuIlePheAlaPhePheGlyValAlaIleLeuSerLeuCyseGly 20
 DB 121 ATGAGGTCCTTTCTCTCTCTGCTGCTGTTCACTTTCTCTGCGCTCTGGA 180
 QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
 DB 181 AAGGTAATATACAGAGGTGTCTTCTCAGCGAACATTTCAAGAAATAAAGAAATATA 240
 QY 41 AlaSerCySGlyAspValAlaIleLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
 DB 241 GCCAATATGAAGATCTTCTGCTTAACATTTCACTTCTGTTATGGAATAACAG 300
 QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
 DB 301 AACCGGTGATATGACGTTTGGGACTTCTAGTTGATCTGTGGACCCAGACTGAGTGGC 360
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnGlnAspGlyLeu 100
 DB 361 TCTAAGAACTAGAGAAAGCTATCCAAATCATGATCAAAACCTGCAACAGATGGCTG 420
 QY 101 GluLysValAlaIleLeuGluProValArgIleProHisTPrGluArgGlyGluIleSerAla 120
 DB 421 GAAGACGTCACCTGAGACAGCTCAGATCACTCTCTGGGGAGGGGCGAAGAACTCTCA 480
 QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
 DB 481 GTGATGTGTGTGCTCGAATTCACAAAGTTGGCTATTATTTAGGCTTGGCGGACGATGGG 540

QY 141 ThrProGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160
 DB 541 ACTCTCTGTAAGGTATCACAGCAAGATCTCGTGGCTCTTTTGTGAACTTCAA 600
 QY 161 ArgArgLAserGluAlaArgGlyLysIleValAlaTyrAsnGlnProTyrIleAsnTyr 180
 DB 601 AGAAGGCATCAGAGGCAAGGGAGAGATTGTTGTATATACAGCTTACATCACTAT 660
 QY 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaIleLysValGlyAlaLeu 200
 DB 661 GGGAAACTGCGCATGTCGGGAGCCGGAGCTGTGAACCTGCAAGGTGGGGCGGTG 720
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisTPrGlyIleGln 220
 DB 721 GCATCCCTCATCCATAGTATGATGCTCTTTTCCATCTACAGCTCTTCAACAGCTCATCA 780
 QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 DB 781 GGATATCAAGATGTGTGCCCAAGATTCCAACAGCTGTATCCAAATATGAAGATGCAGAA 840
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 DB 841 ATGATGTCTCGAATGCGCTTCTCGTGGGACAAAATTGTCATTCATCGAATAATGGAGCA 900
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
 DB 901 AAGGCTTATCCAGATATAGATTTCTTCAACACTTCTTCAAGATCATCTGGAGAGCAATAT 960
 QY 281 ProGluGlnValAlaLeuValSerGlyAlaLeuAspSerTPrAspValGlyGlnGlyAla 300
 DB 961 CCAGAGAGATGTGCTGTCATGAGACATCTGGACGTGGGACGTGGGGACGGTGCA 1020
 QY 301 MetAspAspGlyGlyValAlaPheIleSerTyrGlnAlaLeuSerLeuIleLysAspLeu 320
 DB 1021 CTGATATGCGCGTGGAGCTTCATATCATGGAACACCTCTCATCTTGTAAAGACTT 1080
 QY 321 GlyLeuArgProLysArgThrLeuArgGluValLeuThrAlaGluGlnGlnGly 340
 DB 1081 GGCTGCTGTCMAAGAGACTCTGCTGCTGTGCTCTGACCCGCAAGAAACAGAGGG 1140
 QY 341 ValGluAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
 DB 1141 GTTGTGCTCCCGCATATATATGAGCTATAGGCAATATTTTCCAGATGACAGTTGGTG 1200
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrArgLysSerGlyLys 380
 DB 1201 ATGAGGCTGATCTCAGGAACCTTCTTACCACTGGGCTGACAGTTCACTGGCAGTGAACAG 1260
 QY 381 AlaArgAlaIleMetGluGluValAlaMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 DB 1261 GCCAGGCTTATCATGAAGAGATCATAGTCTCTCTGACCCCTCATATACACCAAGTCT 1320
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTPrIleGlnAlaGlyValProGlyAla 420
 DB 1321 TTTAATGATCAGAGAGAACTGACATTAATTCTGATGCCAAGTGGAGCTCGAGGCC 1380
 QY 421 SerLeuLeuAspPheLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
 DB 1381 AGTCTGGAATATGACTTGTACAGTATTTCTTTTCCATCATATCCCTGAGACACCATG 1440
 QY 441 ThrValMetAspProLySGlnMetAsnValAlaAlaIleValTPrAlaValIleSerTyr 460
 DB 1441 ACTGCCATGATCCAAAGCAAGATGAATGTTGCTGCTCTGTTGGGCTGTGTGCTTAC 1500
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
 DB 1501 GTTGTGCAGACATGAGGAAATGCTGCCAGGTCC 1536
 RESULT 13
 AF097723
 LOCUS AF097723 1778 bp mRNA linear ROD 10-NOV-1998
 DEFINITION Rattus norvegicus hematopoietic lineage switch 2 related protein

[illegible]

QY	101	GIuYrVaIhIshLeuGInPProVaIArgIeIProhIstRgIuRgGIuYgInIuSeRAla	120
Db	421	GAAAGCTCCAGCTCGAGCAGCTCAAGATACCTCACTGGGGAGGGGCGAAGATCTGCA	480
QY	121	ValMetLeuGInPProArgIIehIshYsIleAlaIleLeuGIleuGIYSerSerIeGIY	140
Db	481	GTGATGGTGGGCTCCGAATTCAACAAATTGGCTATTTTAAAGCCTTGCGCAGCATTTGGG	540
QY	141	ThRProGInGInYIIeThRAlaGInValLeuValThRserPheAspGIuLeuGIn	160
Db	541	ACTCTCTCGAAGGATACACAGCAAGAGTCTGGTGGCTCTTTTGTTGAACITCA	600
QY	161	ArgArgAlaSerGIuAlaArgGIYIshIleValValYrAsnGInPProYrIIeAsnTYr	180
Db	601	AGAGGGGATCAGAGGCAAGGGAAGATTGTTTATTAACCAAGCCTTACACTGACTAT	660
QY	181	SerArgThRValGIuTYrArgThRGIuValAlaGIuAlaAlaIshYsValGIYAlaLeu	200
Db	661	GGGAAACCTGGCAGTACCGGAGCGCGAGCTGTGAACCTGGCAAGTGtGGGGCCGTG	720
QY	201	AlaSerLeuIleAsGSerValAlaSerPheSerIleTYrSerPProIstRnGIleGIn	220
Db	721	GCATTCCTCATCCGATCAGTACGTACTTTTCCATCTACAGTCTTCAACAGGTCA	780
QY	221	GIuTYrGIuAspGIYAlaIProYsIleProThRAlaCysIleThRValGIuAspAlaGIu	240
Db	781	GGATATCAAGATGGTGTGCCCAAGATTCCAACAGCCTGTATCACAATAGAAGATGCAGAA	840
QY	241	MetMetSerArgMetAlaSerHisGIYIleYsIleValIleGInLeuYrMetGIYAla	260
Db	841	ATGATGTCTCAATAGGCTTCTCTGGTGGACAAATTTGCATTCATCTGAAATGGGAGCA	900
QY	261	LYeThTYrProAspThRAspSerPheAsnThRValaGIuIleThRGIYSerIeTYr	280
Db	901	AAGACTATCCAGATACAGATTCTTCTCAACACTGTTCAGAGATCACTGGGAGCAAGTAT	960
QY	281	ProGIuGInValIleValLeuValSerGIYHisLeuAspSerTrpAspValGIYGIuGIYAla	300
Db	961	CCAAAGGAGAGTGTCTCGGTCAAGGACATCTGGACAGCTGGGACGTGGGCGAGGTGCA	1020
QY	301	MetAspAspGIYGIYGIYAlaPheIleSerTrpGIuAlaLeuSerLeuIleYAspLeu	320
Db	1021	CTGATGATGGCGGTGAGCCTTCAATTCATGGAGACCTCTCACTTTGTTAAGATCTT	1080
QY	321	GIYLeuArgPProYsArgThRLeuArgLeuValLeuTrpThRAlaGIuGIuGIuGIYGIY	340
Db	1081	GGGCTGCCTCCAAAGAGGACTCTCGCTGGTGGTCTTGACCGCAAGAAACAAGGAGG	1140
QY	341	ValGIYAlaPheGInTYrTYrGIuLeuHisIshYsValAsnIleSerAsnTYrSerLeuVal	360
Db	1141	GTTGGTGTCTCCAGTATATGACTCAATDAAGCAAAATTTTCCAAGTACAGTTGGTG	1200
QY	361	MetGInSerAspAlaGIYThRPhLeuProThRGIYLeuGInPheThRGIYSerGIYIshYs	380
Db	1201	ATTGAGCGCTGACTCAGAAACCTTTTACCACTGGCTGCAGTCAACCGGAGTGCACAG	1260
QY	381	AlaArgAlaIleMetGIuGIuValMetSerIeLeuGInPProLeuAsnIIeThRGIuVal	400
Db	1261	GCCAGGGCTATCAAGAGAGATAGTATGCTCTCGCAACCCCTCAATATACCAAGAGCT	1320
QY	401	LeuSerHisGIYGIuGIYThRAspIIeAsnPheTrpIleGInaGIYValProGIYAla	420
Db	1321	TTTATATGTCAGAAAGAACTGACATTTACTTCTGATCCAAAGTGGAGTGCCTGGAGCC	1380
QY	421	SerLeuLeuAspAspLeuTYrLYeTYrPhePhePheHisIshSerHisGIYAspThMet	440
Db	1381	AGTCTGCAGATGACTTGACAAAGTATTTCTTTTCCATCTATCCCAATGAGACACCATG	1440
QY	441	ThRValMetAspPProYsGIuMetAsnValaIalaIalaValITrPAlaValIleValSerTYr	460
Db	1441	ACTGCCATGAGATCCAAAGCAAGATAAATGTGCTCTGTGTGGCTGTGTGTGCCTTAC	1500

QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
 DB 1501 GTTGTGCAGACATGAGAAATGCTGCCAGATCC 1536
 RESULT 14
 AF009513
 LOCUS AF009513 1716 bp mRNA linear ROD 01-JUN-1998
 DEFINITION Mus musculus hematopoietic lineage switch 2 (HLS2) mRNA, complete cds.
 ACCESSION AF009513
 VERSION AF009513.1 GI:3169728
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1716)
 AUTHORS Williams, V.H., Chan, C.-Y. and Klinken, S.P.
 TITLE Hematopoietic Lineage Switch 2 (HLS2), a novel mRNA species induced during an erythroid to myeloid lineage switch
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1716)
 AUTHORS Williams, V.H., Chan, C.-Y. and Klinken, S.P.
 TITLE Submitted (20-JUN-1997) Biochemistry Department (Laboratory of Cancer Medicine), University of Western Australia, Royal Perth Hospital, Wellington Street, Perth, WA 6001, Australia
 JOURNAL
 FEATURES
 source
 1..1716
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1..1716
 /gene="HLS2"
 97..1398
 /gene="HLS2"
 /codon_start=1
 /product="hematopoietic lineage switch 2"
 /protein_id="AAC17945.1"
 /db_xref="GI:3169729"
 CDS
 1..1716
 /translation="MRSLEPFLFVHLALGSGKAVFKGVSGRTREIKEEIANVEDV AKATINLAVGKYQNRYSVERGLADVTPRISGSKNLEKAIQIWMYQLQDDLENVH LEQVRIPIHMERBESAVMLEPRIHMAILIGLSSIGTPGGITAEVLVASDELQOR ASEARGLIIVNOPYGAEKTYOVQVGAENAKGAVASLIQSVASFISPHGIQ KYODGPKITPTACTVEDABEMMSRMASSGNKIVIHLEMGAKTPTDSFNTVAETTS MYRDEVGGALDDGGAGARTISWALSIVYDGLGRPKRTIRLVMTREBGGCGASGYE LHKAVISKISLVMEADSGTFLEPTGLQFTGSDAKAIMEKVMMLQPLNVTIKVFSVGG TDINFWIQAGVGAELRDLDLYKFFPHSHSGDTMTSWIQSR"
 BASE COUNT 476 a 370 c 414 g 456 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.06e-146 Length: 1716
 Score: 2029.00 Matches: 402
 Percent Similarity: 90.89% Conservative: 27
 Best Local Similarity: 85.17% Mismatches: 29
 Query Match: 83.95% Indels: 15
 DB: 10 Gaps: 2
 US-09-745-763-36 (1-472) x AF009513 (1-1716)
 QY 1 MetLysPheLeuIlePheAlaPheGlyValHisLeuLeuSerLeuCySerGly 20
 DB 97 ATGAGTCCCTTTCTTCTCTGTTCACT-----GTTCACTTTTACCTCTGGCTCTGCT 150
 QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGlu 40
 DB 151 AAAGCTGATTCAGAAATGCTGTTCTCAGCAACATTCGAGAAATATAAAGAAATA 210
 QY 41 AlaSerCyGlyArgValAlaLysAlaIleIleAsnLeuAlaValIleGlyLysAlaGln 60
 DB 211 GCCAATCTAGAAATGCTGCTAAAGCAATATCAACCTTGCTTTATGTAATAATACAG 270

QY 61 AsnArgSerIleGlyArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
 DB 271 AACGGTCTTATGAGCGCTTTGGACTTCTTATGTAATCTGTTGGACCGACGATGGTGC 330
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetCysTrpGlnAsnLeuGlnAspGlyLeu 100
 DB 331 TCTAAGAACCTTAGAAAGCAAGCTATTCATCTGTACCAAAACCTCGACGAGCAATGGGCTA 390
 QY 101 GluLysValHisLeuGluProValArgIleProHisIleTrpGluArgGlyGluGluSerAla 120
 DB 391 GAAATGTTCACTCGACGACAGCTCAGAAATACCCACTGGAGAGGAGGAGAAATCTGCA 450
 QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
 DB 451 GTGATGCTTGAGCGCTCGAATTCACAAATGCTATTGCTTGGACGACGATTCGG 510
 QY 141 ThrProGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160
 DB 511 ACTCCCGCAGAGGACATCACAGCAAGTGTGGTGTAGCCCTCTTTGATGAACCTTCA 570
 QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
 DB 571 AGAAGAGCATCAGAAAGCAAGGAGATCATTTGTTATTAACAGCTTACATCGCTTAT 630
 QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200
 DB 631 GAGAAAGCTGCGCATGCCGCGGTCCAGGAGACTGTGAGAGCTGCCAAGTGGAGCTGTG 690
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleGlnIleGln 220
 DB 691 GCATCCCTCATCATCATCATGATGCTTCCTTCCATCTACAGTCCCTCACAGGATTCAG 750
 QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
 DB 751 AAATATCAAGATGGTGTGCCCAAGATTCACACCTGTATCAACAGTACAGAAATGCGCAA 810
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
 DB 811 ATGATGCTCGAATGGCTTCGATCGGGAACAAATTTGCATTCATTCGAAATGGAGCA 870
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
 DB 871 AAGACATACCAAGATACGATCTTCATCATCTGTGAGAGATCAGGAGCATGATAT 930
 QY 281 ProGluGlnValLeuLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
 DB 931 CCAAG-----GAGTTGGCAGGGTGCA 954
 QY 301 MetAspAspGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
 DB 955 CTGATGATGCTGATGGAGCTTCATATCATGGGAGACCTCTCACTTTGTAAGATCTT 1014
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGlyGly 340
 DB 1015 GGGCTGCTCCAAAGAGAGACTCTCGCTGCTGTGAGATGCGAGAAACAGGAGGA 1074
 QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
 DB 1075 ATTGTGCTCCCAATATGATGCTCATTAAGCAATATTTCCAAATACAGCTGTG 1134
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheHisGlySerGlyLys 380
 DB 1135 ATGAGAGCTGATCAGAGAACTTCTTACCACTGAGCTGAGCTTCACTGCGAGAGACAAG 1194
 QY 381 AlaArgAlaIleMetGluGlnValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
 DB 1195 GCCAGGCTATCAGAAAGAGAGTATCATGATCTTGTGCAACCCCAATGTCACCAAGGTC 1254
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
 DB 1255 TTTATGATAGGAGAGAACTGATATTAATCTTGATCCCAAGCTGAGAGTCCGAGACC 1314
 QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440

Db 1315 AGTCGCGTGAATGATCTTGACAGATTTCTTTTCCATCATCCACGAGACCATG 1374
Qy 441 ThrValMetAspPolysGluMetAsnValAlaIaIaValTPAlaValSerTyr 460
Db 1375 ACGTC-ATGGATCCAAAGACAGATGATGTTCTGTGTATGAGCGCTGTGCTTAT 1433
Qy 461 ValValAlaAspMetGluGluMetLeuProAspSer 472
Db 1434 GTTGTTCAGACATGATGATGATGCTGCCAGGTCC 1469
RESULT 15
BD059610 609 bp DNA linear PAT 27-AUG-2002
LOCUS Secreted expressed sequence tags (ESTs).
DEFINITION BD059610
ACCESSION BD059610 GI:22605216
VERSION BD059610.1 GI:22605216
KEYWORDS JP 2001519666-A/1465.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 609)
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
Treacy,M., Spaulding,V. and Agostino,M.J.
TITLE Secreted expressed sequence tags (ESTs)
JOURNAL Patent: JP 2001519666-A 1465 23-OCT-2001;
GENETICS INSTITUTE INC
COMMENT PN JP 2001519666-A/1465
PD 23-OCT-2001
PF 10-APR-1998 JP 1998543068
PR 10-APR-1997 US 08/835913
PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE,PI
DAVID MERBERG,
PI MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC Strandedness:
Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source Location/Qualifiers
1..609
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
BASE COUNT 190 a 127 c 148 g 144 t
ORIGIN
Alignment Scores:
Pred. No.: 6.01e-58 Length: 609
Score: 870.00 Matches: 176
Percent Similarity: 99.44% Conservative: 0
Best Local Similarity: 99.44% Mismatches: 1
Query Match: 36.00% Indels: 1
DB: 6 Gaps: 0
US-09-745-763-36 (1-472) x BD059610 (1-609)
Qy 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20
Db 71 ATGAATATCTCTTATCTTCGATTTTTCGATGTTTCACTTTATCCCTGAGCTGCGG 130
Qy 21 LysAlaIleCysValAsnGlyIleSerLysArgThrPheGluGluIleLeuGluIle 40
Db 131 AAAGCTATATGCAAGAAATGCAATCTCTTAAGAGACTTTTGAAGAAATAAAGAAATA 190
Qy 41 AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
Db 191 GCCACGTGTGAGATGTTGCTTAACCAATCATCACTAGCTTTATGTTAAAGCCAG 250
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db 251 AACGATCTCTATGAGCGATTGGCACTTCTGTTGATACTGTTGAGACCACTGAGTGGC 310

Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
Db 311 TCCAAAGACTA-GAAAAAGCATCCCAATTATGACAAAACCTGCAGCAAGATGGCTG 369
Qy 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluSerAla 120
Db 370 GAGAAAGTCACTGAGCCAGCAGTGAATACCCCACTGGAGAGAGGAGAAATCAGCT 429
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 430 GTGATGCTGAGCCAGAAATTCATAGATGACCATCTCGAGTCTTGACAGCATTTGGG 489
Qy 141 ThrProGluGluIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
Db 490 ACTCCTCAAGAGGCAATTACAGCAGAGATTCTGTGTGATCCTTTGATGAATCAG 549
Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyr 177
Db 550 AGAAGGCGCTCAGAAAGCAGAGAGGAGATTTGTTTATACCAACCTTAC 600

Search completed: December 23, 2003, 01:04:01
Job time : 4382 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)